

UNITED STATES AIR FORCE
SUMMER RESEARCH PROGRAM -- 1995
SUMMER FACULTY RESEARCH PROGRAM FINAL REPORTS

VOLUME 2B
ARMSTRONG LABORATORY

RESEARCH & DEVELOPMENT LABORATORIES
5800 Uplander Way
Culver City, CA 90230-6608

Program Director, RDL
Gary Moore

Program Manager, AFOSR
Major David Hart

Program Manager, RDL
Scott Licoscos

Program Administrator, RDL
Gwendolyn Smith

Program Administrator, RDL
Johnetta Thompson

Submitted to:

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Bolling Air Force Base

Washington, D.C.

December 1995

Reproduced From
Best Available Copy

19981218 060

REPORT DOCUMENTATION PAGE

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for preparing and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations, 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC 20503.

Approved		
AFRL-SR-BL-TR-98-		
0825		
1. AGENCY USE ONLY (Leave Blank)		2. REPORT DATE December, 1995
3. REPORT TYPE Final		
4. TITLE AND SUBTITLE USAF Summer Research Program - 1995 Summer Faculty Research Program Final Reports, Volume 2B, Armstrong Laboratory		5. FUNDING NUMBERS
6. AUTHORS Gary Moore		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Research and Development Labs, Culver City, CA		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) AFOSR/NI 4040 Fairfax Dr, Suite 500 Arlington, VA 22203-1613		10. SPONSORING/MONITORING AGENCY REPORT NUMBER
11. SUPPLEMENTARY NOTES Contract Number: F49620-93-C-0063		
12a. DISTRIBUTION AVAILABILITY STATEMENT Approved for Public Release		12b. DISTRIBUTION CODE
13. ABSTRACT (Maximum 200 words) The United States Air Force Summer Faculty Research Program (USAF- SFRP) is designed to introduce university, college, and technical institute faculty members to Air Force research. This is accomplished by the faculty members being selected on a nationally advertised competitive basis during the summer intersession period to perform research at Air Force Research Laboratory Technical Directorates and Air Force Air Logistics Centers. Each participant provided a report of their research, and these reports are consolidated into this annual report.		
14. SUBJECT TERMS AIR FORCE RESEARCH, AIR FORCE, ENGINEERING, LABORATORIES, REPORTS, SUMMER, UNIVERSITIES		15. NUMBER OF PAGES
		16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified
		20. LIMITATION OF ABSTRACT UL

PREFACE

Reports in this volume are numbered consecutively beginning with number 1. Each report is paginated with the report number followed by consecutive page numbers, e.g., 1-1, 1-2, 1-3; 2-1, 2-2, 2-3.

Due to its length, Volume 2 is bound in two parts, 2A and 2B. Volume 2A contains #1-23, and Volume 2B contains reports #24-43. The Table of Contents for Volume 2 is included in both parts.

This document is one of a set of 16 volumes describing the 1995 AFOSR Summer Research Program. The following volumes comprise the set:

<u>VOLUME</u>	<u>TITLE</u>
1	Program Management Report <i>Summer Faculty Research Program (SFRP) Reports</i>
2A & 2B	Armstrong Laboratory
3A & 3B	Phillips Laboratory
4	Rome Laboratory
5A, 5B, & 5C	Wright Laboratory
6A & 6B	Arnold Engineering Development Center, Wilford Hall Medical Center and Air Logistics Centers <i>Graduate Student Research Program (GSRP) Reports</i>
7A & 7B	Armstrong Laboratory
8	Phillips Laboratory
9	Rome Laboratory
10A & 10B	Wright Laboratory
11	Arnold Engineering Development Center, Wilford Hall Medical Center and Air Logistics Centers <i>High School Apprenticeship Program (HSAP) Reports</i>
12A & 12B	Armstrong Laboratory
13	Phillips Laboratory
14	Rome Laboratory
15A&15B	Wright Laboratory
16	Arnold Engineering Development Center

SFRP FINAL REPORT TABLE OF CONTENTS**i-xiv**

1. INTRODUCTION	1
2. PARTICIPATION IN THE SUMMER RESEARCH PROGRAM	2
3. RECRUITING AND SELECTION	3
4. SITE VISITS	4
5. HBCU/MI PARTICIPATION	4
6. SRP FUNDING SOURCES	5
7. COMPENSATION FOR PARTICIPATIONS	5
8. CONTENTS OF THE 1995 REPORT	6

APPENDICIES:

A. PROGRAM STATISTICAL SUMMARY	A-1
B. SRP EVALUATION RESPONSES	B-1

SFRP FINAL REPORTS

CREATING ADAPTABLE INFORMATION DELIVERY SYSTEMS FOR TEAM-BASED
DISTRIBUTED DECISION MAKING DURING SUSTAINED OPERATIONS:
THE APPLICATION OF TECHNOLOGICAL COUNTERMEASURES

Robert P. Mahan
Associate Professor
Department of Psychology

University of Georgia
Athens, Georgia 30602

Final Report for:
Summer Faculty Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Armstrong Laboratory

August 1995

CREATING ADAPTABLE INFORMATION DELIVERY SYSTEMS FOR TEAM-BASED
DISTRIBUTED DECISION MAKING DURING SUSTAINED OPERATIONS:
THE APPLICATION OF TECHNOLOGICAL COUNTERMEASURES

Robert P. Mahan
Associate Professor
Department of Psychology
University of Georgia

Abstract

The exponential growth of flexible automated systems is transforming many of the tasks performed by air crew operators into command and control operations. We are seeing air crew teams indirectly controlling large tactical systems via their interaction with computers which now directly control system parameters. Operators must interact with highly flexible technology which provides numerous functions and options for carrying out tasks under different situations. Further, operators must execute these activities under very demanding operational circumstances. Distributed situational awareness in the state of avionic systems embedded in sustained combat scenarios will become a crucial component in mission outcomes, as well as in safety compliance activities, and the avoidance of serious operational hazards. This report serves to introduce a network of ideas that will lead to the development of adaptable systems that can accommodate the complexity of command and control tasks and the psychophysiological state changes that occur during sustained duty. An extended bibliography that provides an extensive literature index is included in the present report.

CREATING ADAPTABLE INFORMATION DELIVERY SYSTEMS FOR TEAM-BASED
DISTRIBUTED DECISION MAKING DURING SUSTAINED OPERATIONS:
THE APPLICATION OF TECHNOLOGICAL COUNTERMEASURES

Robert P. Mahan

Introduction

Performance in complex systems relies on group or team functioning rather than the isolated behavior of a single individual. Thus, it is not surprising that a great deal of research has focused on decision making at the team level. A premise that has been developed elsewhere and which we continue to develop here is that teams are composed of many different kinds of individuals with many different types of expertise. Thus, task efforts that combine expertise in order to achieve a level of awareness of a particular state of the world may best be represented as a hierarchical team judgment problem. In aviation systems there are many kinds of expertise from navigation, communication, surveillance, to threat identification, maneuver, and engagement. Here judgment is a distributed process that relies on the execution of expertise by different people who possess knowledge about different parts of avionic systems. Brehmer and Hagafors (1986) were among the first to formalize a quantitative description of hierarchical judgment at the individual level and Hollenbeck et al (1995) have since expanded these notions to include team level decisions.

The basic principle behind developing displays relies on the idea that complex information can be displayed in a manner that can be quickly understood with very little mental effort (i.e., mental workload) necessary for this understanding. Thus, the goal is to select a display format where the meaning of the information presented is highly evident and does not require a detailed and intensive off-line analysis. However, prior to describing the physical characteristics of such a display it is necessary to document what kind of cognitive processing mode we are after. That is, how do we want the air crews to think about the information that is contained in the display. More precisely, what mode of cognition is most compatible with the nature of global (holistic) assessments of avionic systems. As we shall see, knowing the answer to this question will determine, at least in

part, the physical properties of the display that we construct in order to support this kind of cognitive behavior.

Relevant Cognitive Science

Intuitive versus Analytical Cognitive Functioning

Intuitive cognition is often presumed to be distinguishable from analytically-based modes of cognition in the sense that intuition tends to be a kind of seamless form of cognition that manifests parallel processing characteristics (i.e., quick, effortless, single stage). Kahneman, Slovic and Tversky (1982) characterize it as an informal reasoning process that lacks the application of formal methods of calculation. Most researchers agree that intuitive cognition is fundamentally inferential, where, for example, several propositions regarding relationships about a given phenomenon are combined with general knowledge to yield a verbal conclusion (Johnson-Laird, 1990; Evans, 1990). Some have been more poignant in differentiating between these forms of cognition. Johnson-Laird (1990) notes when characterizing the distinction between analysis and inference, -" it is one thing to multiply two numbers, and quite another to draw a picture, to compose a sonata, or to write a poem" (p. 156). Further, he goes on to suggest that inference may be associated with the same cognitive mechanisms that are brought to bear when one "uses one's imagination". Although, intuitive cognition has typically been difficult to operationalize, it clearly has been conceptually differentiated in the past from analytical modes of cognition (e.g., Beach and Mitchell, 1978; Brooks, 1978; Von Winterfeldt and Edwards, 1986; Kahneman, Slovic and Tversky, 1982; Garner, 1981).

In contrast to intuitive cognition, analytical cognition is typically conceived as being a highly proceduralized and deductive serial process that conforms to the basic cannons of a particular system of logic. Confirmation to the cannons of logic insure the validity of solutions. Hammond (1981) notes that the logic system serving as the framework for supporting analytical cognition can be; "a) mathematical in nature (e.g., simultaneous equations in word algebra problems), b) statistical (e.g., Bayesian or Fisherian statistics for uncertain events), c) propositional logic (e.g., modus tollens, modus ponnens), d) problem solving strategies (such as opening and end games in chess), e) scientific laws (as in physics and chemistry) and f) ideological maxims (in political debates)", (pp. 15).

Researchers typically acknowledge (a) intuition is a mode of cognition that is used daily by people in natural decision making environments and (b) that it is qualitatively different from analytically driven problem solving. The majority of attempts at elaborating on the distinction between analytic and intuitive forms of cognition have done so by illustrating the profound shortfalls inherent in relying on intuitive problem solving skills in decision making. Tversky and Kahneman (1974) and later Kahneman, Slovic and Tversky (1982) present an extensive review of the errors and biases related to intuitive cognition that take the form of heuristic strategies in decision making. However, very few studies have adequately defined what intuition is, nor how an individual's intuitive processing compares with his/her analytical skills. The latter point is best dramatized by Hammond, Hamm, Grassia and Pearson (1987) who cogently dispute that virtually all major research efforts have benchmarked a person's skill at exercising intuitive judgment with an analytically derived algorithm or axiomatically defined rule system. They have stressed that using a prescriptive model criterion index is unsatisfactory, and at best, an indirect method of characterizing the efficacy of intuitive cognition. Some have taken a different (albeit provocative) approach in disputing the shortfalls in complex human decision making. For example, Anderson (1986) argues that many of the notorious and alleged decision biases reported in decision making research are simply manifestations of poor experimental/methodological techniques, fallacious experimental interpretations and artifactual laboratory findings.

Figure 1 illustrates a model that distinguishes between an object or condition that is defined by various information sources (cues), and the psychological representation of the object or condition which is defined through a particular judgment policy. The lens model portrays the environment as a series of cues whose relationships with the environment are less than perfect. A decision maker is viewed as interacting with his or her environment through a 'lens' which is often distorted because of this imperfect and uncertain relationship. The relationship between the cues and the environment is typically characterized by "ecological validities" that, in theory, can range in absolute value from 0 to 1.0. Ecological validity represents the predictive importance of each cue.

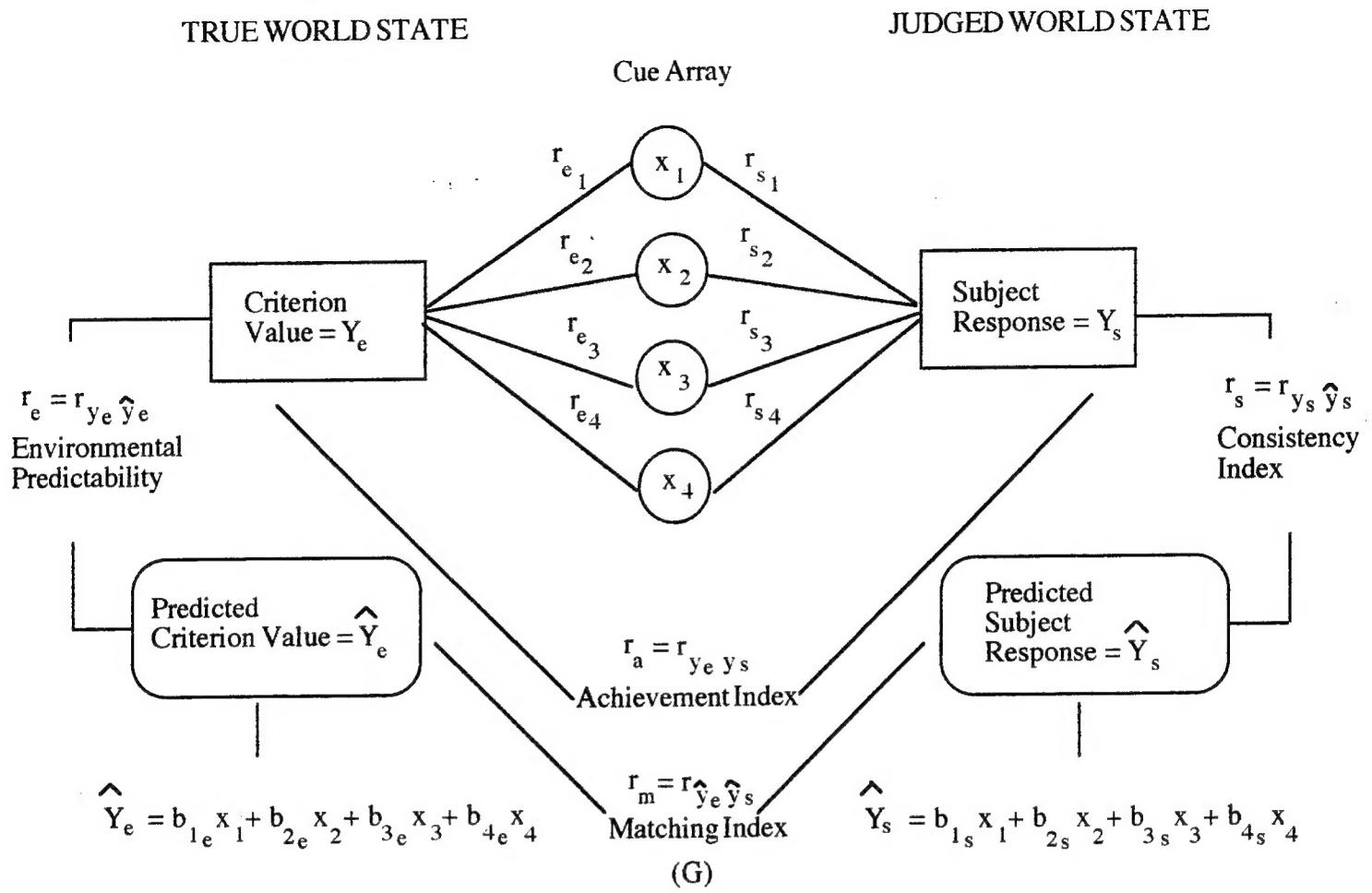


Figure 1. Lens model of judge and environment interaction.

Discussion on Implications of Cognition for Display Design

The implications of the above discussion on display design are simple and direct.

Performance by experts (as well as non-experts to a lesser degree) depends on the situation in which they work, and the task requirements for performance. That is, tasks often call for both analytical and intuitive processing. In this case the task would induce a quasi-rational mode of cognition that lies somewhere in between the end points on a cognitive continuum.

An additional implication of the above discussion is the notion of information format. If one accepts, at least in part, that the task is responsible for structuring and inducing a particular form of cognition, then it follows that the structure of the task must be preserved in the physical display itself. This means that the transformation from the task to the

display of task parameters must be invariant so as to induce the appropriate organizing principle called for by the structural characteristics of the task. From this perspective, it is clearly possible to have a task calling for a particular cognitive mode, yet this mode is not supported by the physical properties of the display interface. This disconnect between task structure and display structure will likely lead to ineffective cognitive performance. The physical properties of the display itself are crucial for preserving task structure.

Operational Context: Sustained Duty

A fundamental component of operational readiness in military doctrine and training is preparing soldiers for sustained work that extends beyond normal duty hours. Command and control, medical, security, communications, navigation, and most transportation activities demand continuous and often sustained work. Sustained operations (susops) are typically differentiated from continuous operations (conops) in that while conops is a round-the-clock operational environment, there is down-time or periods of very low activity where operators can secure rest. In sustained operations, the work is unceasing. Further, susops is a characteristic of tactical combat or high-alert conditions and is often defined by surge-states (increased tempo of performance). As a result of the requirement for sustained vigilance and performance during susops, individuals and teams do not have adequate opportunities for rest, making sleep-loss and fatigue highly influential factors in mission outcomes. It is noteworthy that sustained performance variables have been identified as being associated with the underlying causes of many industrial and military accidents (Office of Technology Assessment, OTA-BA-463, 1991).

Information on how sustained operations affect many command and control decision activities remains largely unknown. The absence of research in this area is especially disturbing considering that the exponential growth of flexible automated systems is essentially transforming all tasks performed by all human operators into command and control tasks. We are seeing humans indirectly controlling large tactical systems via their interaction with computers. It is the computers which now directly control system parameters. This means that the human operators are responsible for interacting with highly flexible technology that provides large numbers of functions and options for carrying out tasks under different situations. Operators are now required to integrate large

amounts of information in selecting system modes best suited for a particular situational circumstance. This means that operators must have a reasonably clear understanding of the "situational state" prior to selecting a particular mode of operation. The task requirements necessary for human functioning in flexible automated systems is calling on much greater levels of information integration activities, and thus are similar to those seen in command and control type decision making behaviors. Even our secretarial staff is faced with situational awareness problems inherent in flexible technology, when during word processing they must determine - "what mode am I in and how in the world did I get there?".

The Absence of Theory

We still have little in the way of a theoretical framework that allows predictions of cognitive behavior in a variety of tasks performed during sustained operations. One might assume that the stress literature may help our case since its reasonable to classify susops as a stressor. However, as many have noted, the stress literature is currently a malgramation of hundreds of studies which point in hundreds of directions. This is largely due to the absence of a unifying theoretical framework that would serve to direct questions asked about stressors and stress responses. Some researchers have advocated striking the word "stress" from scientific parlance altogether because it is so misused and it does not discriminate among task states or human conditions. Those same people favor using specific terms like "time-on-task" or "time-of-day" as the stress component which makes the operational parameter of interest explicit. Thus, standard contexts for measuring behavior can be created (unlike the scenario that exists today where stress can mean almost anything).

Recent evidence suggests that temporal factors, such as continuous performance, differentially affect analytical versus intuitive cognition (Mahan, 1991b, 1992b, 1994). For example, it appears that the task duration decrement for an intuitive task tends to be associated with a reduction in precision of performance, as opposed to rate of performance, which has been a traditional finding in continuous performance research employing analytical tasks (Mahan, 1991b). Further, the performance decrement for intuitive tasks seems to manifest itself as a loss in the ability to consistently apply personal judgments

regarding some criterion variable. Decrements in judgment consistency have been viewed in the past as reductions in the ability to control the execution of a learned policy for making intuitive judgments (Hammond & Summers, 1972; Hammond & Wascoe, 1980).

Hammond and Summers (1972) further argue that controlling knowledge execution is independent of a subject's over-all task knowledge state. That is, one may intellectually understand the necessary requirements for task performance, yet be unable to actively control the application of that knowledge when actually performing a task.

Implications of Operational Context on Situational Assessment

The findings in the above work as well as recent studies (see, Mahan 1994 for review) support the utility in using the cognitive continuum theory as a framework in which to evaluate the effects of continuous and sustained operations on complex decision making. Overall, the data appears to indicate that the ability to make multidimensional judgments is affected by both task duration and task uncertainty. The primary fatigue-induced impairment tends to be associated with the individual's control of the execution of knowledge regarding the judgment task.

Cognitive Engineering: The Delivery System

Display Theory Background

As we have seen, information processing within complex decision environments has elicited the work conducted in perception (e.g. Brunswikian views, 1956). A large amount of research in perception exists on the perceptual processing of multidimensional stimuli. Historically the interest in this area has focused on the characteristics of the processing system per se. However, more recent attention has been placed on investigations in the inherent structural properties of stimuli and their effect on the perceptual process. Garner (1970) stimulated interest in this area by arguing that before one can understand the details of perceptual processing, one must understand the details associated with the structure of the stimulus.

A central issue in the research on stimulus structure has been the nature of the relation between dimensions which represent a multidimensional stimulus. Garner (1974) has discussed two major ways in which stimulus dimensions can be related. Stimuli composed of integral dimensions produce a Euclidean metric in direct distance scaling, facilitate the

discrimination of stimuli on one dimension when another dimension varies in a correlated manner, and inhibit the discrimination of stimuli on one dimension when another dimension varies in an orthogonal manner. An example of integral dimensions would be the hue and saturation of a color (Garner, 1970). Separable dimensions (e.g., separate vertical bars) produce a city-block metric in direct distance scaling and produce neither facilitation with correlated dimensions nor interference with orthogonal dimensions in a discrimination task. Psychologically, integral dimensions appear as an integrated whole, whereas separable dimensions are seen as distinct and separate. Further, dimensions tend to be integral if the existence of one dimension depends on the existence of another dimension. For example, any one of the dimensions of color (i.e., brightness, hue, and saturation) can not exist without values on the other dimensions.

The study of integral and separable dimensions has traditionally been limited to relatively simple stimuli and elementary perceptual processes. Stimuli are generally composed of physical attributes such as color or geometric form and are often limited to two binary dimensions. In discrimination and identification studies, the task is to evaluate stimuli on the basis of one dimension while the stimuli vary along another dimension. The rule relating stimuli to responses is based on physical attributes of one of the stimulus dimensions. Since these tasks are generally easy to perform, speeded responses are obtained and reaction time serves as the primary dependent variable. Classification and similarity scaling do not specify what aspects of the stimuli the subject should use in forming classes and assignment ratings. Instead, the interest is on identifying stimulus structure by the way in which the subjects globally view the stimuli. With integral dimensions, classification is based on the overall similarity structure of the stimuli; with separable dimensions, subjects group stimuli on the basis of a single dimension.

Integral dimensions appear to show an increase in the speed of processing when two dimensions are correlated, and interference when the two dimensions are orthogonal (Garner, 1974; Foard and Kemler, 1989). Further, selective attention to a single dimension of an integral stimuli appears difficult to achieve because perception is dominated by the overall similarity structure, or the emergent features corresponding to holistic processing (Smith and Kemler, 1978; Garner and Felfody, 1970). In contrast,

separable dimensions are those that do not show any improvement in the speed of processing when the dimensions are correlated, nor interference when they vary orthogonally. In addition, separable dimensions support focused attention on specific features of the stimulus (Garner, 1974, 1976; Pomerantz, 1981).

Taken together, the above research indicates that the structure of the stimulus plays an important role, not only in elementary perceptual operations, but also in higher order cognitive processes. Specifically, stimuli composed of integral dimensions appear to facilitate performance across several types of perceptual and cognitive tasks. Recent interest has focused on the role of integral dimensions for facilitating performance on information integration tasks.

Discussion on Implications of Display Design for Cognitive Functioning

The above discussion on display configuration implies that there exists a very close link between the judgment process and the structural properties of the display for influencing this process. Oversimplifying the very brief discussion on integral and separable displays, if the task is one of making judgments about a very complex system which is composed of correlated information dimensions, an integral display will be most compatible with the task at hand and will support the mode of cognition most compatible with task structure. In contrast, if the problem domain is well defined and composed of orthogonal information dimensions, a separable display may be more appropriate. Clearly, selecting the wrong display for the task will be counterproductive in the sense that the aviation community will not find it useful. The worse case scenario, of course, is that an inappropriate display will lead to wrong kinds of feedback and the risk of poorly prepared and trained soldiers.

Significance to the Air Force

Improvements in the performance of command and control activities in global and tactical air operations will need to address issues associated with sustained operational contexts. The next generation of countermeasures to combat human performance decrements will use technology as a means to assist and support air crew teams with command and control tasks that must be performed in difficult work environments. Smart, adaptable systems are likely to be the next step toward maintaining human performance in air and airland battle operations.

Extended References

Alluisi, E. A. & Morgan, B. B. (1982) Temporal factors in human performance. In: E. A. Alluisi & E. A. Fleishman (Eds.), Human Performance and Productivity: Stress and Performance Effectiveness, New Jersey: Hillsdale.

Anderson, J. R. (1976). Language, Memory, and Thought. New Jersey: Lawrence Erlbaum.

Anderson, N. H. (1986). A cognitive theory of judgment and decision. In B. Brehmer, H. Jungermann, P., Lourens & G. Sevon (Eds.), New Directions in Research on Decision Making (pp. 63-104). Amsterdam: North-Holland.

Baddeley, A. D. (1968). A three minute reasoning test based on grammatical transformation. Psychonomic Science, 10, 341-342.

Beach, L. R. & Mitchell, T. R. (1978). A contingency model for the selection of decision strategies. Academy of Management Review, 3, 439-449.

Brehmer, B., & Hagafors, R. (1986). Use of experts in complex decision making: A paradigm for the study of staff work. Organizational Behavior and Decision Processes, 38, 181-195.

Brehmer, B., & Joyce, C. R. B. (1988). Human Judgment: The SJT View. New York: North Holland.

Brehmer, B. (1976). Subjects' ability to find the parameters of functional rules in probabilistic inference tasks. Organizational Behavior and Human Performance, 17, 388-397.

Billings, C. E. (1994). Situational awareness in complex systems: A commentary. In R. D. Gilson, D. J., Garland, and J. M. Koonce (Eds.), Situational Awareness in Complex Systems. Daytona: Embry-Riddle Aeronautical University Press.

Broadbent, D. E. (1955). Variations in performance arising from continuous work. In: Conference on Individual Efficiency in Industry. Cambridge, England: Medical Research Council.

Broadbent, D. E. (1971). Decision and Stress. New York: Academic Press.

Brooks, L. (1978). Nonanalytic concept formation and memory for instances. In E. Rosch

and B. Lloyd (Eds.) Cognition and Categorization, New York: Lawrence Erlbaum.

Brunswik, E. (1952). The conceptual framework of psychology. In: International Encyclopedia of Unified Science, Chicago: University of Chicago Press.

Carswell, C. M., & Wickens, C. D. (1987) Information integration and object display: An integration of task demands and display superiority. Ergonomics, 30, 511-527.

Carvalhais, A. B., Tepas, D. I & Mahan, R. P. (1988) Sleep duration of permanent shift workers on workdays and non-workdays. Sleep Research, 17, 109.

Casey, E. J. & Wickens, C. D. (1986). Visual Display Representation of Multidimensional Systems, (Tech. Report CPL -86-2). Urbana-Champaign, University of Illinois, Psychophysiology Laboratory.

Chiles, D. W., Alluisi, E. A. & Adams, O. S. (1968) Work schedules and performance during confinement. Human Factors, 10, 143-196.

Colligan, M. J. & Tepas, D. I. (1986) The stress of hours of work. American Industrial Hygiene Association Journal, 47, 686-695.

Coury, B. G., Boulette, M. D., & Smith, R. (1989). Effect of uncertainty and diagnosticity on classification of multidimensional data with integral and separable displays of system status. Human Factors, 31, 551-570.

Davies, D. R., and Parasuraman, R. (1982). The Psychology of Vigilance. London: Academic Press.

Dement, W. C. (1972) Sleep deprivation and the organization of behavioral states. In: C. D. Clemente, D. P. Purpura & F. E. Mayer (Eds.), Sleep and the Maturing Nervous System, New York: Academic Press.

Dumont, C. (1985) Shiftwork in asian developing countries: An overview. In: Shift Work and Related Issues in Asian Countries, Geneva: International Labor Office Publications.

Dinges, D. F., Orne, M. T., Orne, E. C. & Evans, F. J. (1985) Assessing performance upon abrupt awakening from naps during quasi-continuous operations. Behavioral Research Methods, Instrumentation & Computers, 17, 37-45.

Englund, C. E., Ryman, D. H., Naitoh, P. & Hodgdon, J. A. (1985) Cognitive

performance during successive and sustained physical work episodes. Behavioral Research Methods, Instrumentation & Computers, 17, 75-85.

Evens, J. St. B. T. (1990). Deductive reasoning in human information processing. In: K.A. Mohyeldin Said, W.H. Newton-Smith, R. Viale. and K.V. Wiles (Eds.) Modelling the Mind. Oxford: Clarendon Press, 255-269.

Eyesneck, H. J. (1988). Individual differences, arousal and monotonous work. In: J. P. Leonard (Ed.) Vigilance: Methods, Models, and Regulation. Frankfurt: Lang.

Eyesneck, H. J. (1989). Individual differences in vigilance performance. In: V. Coblenz (Ed.) Vigilance and Performance in Automatized Systems. Boston: Kluwer Academic Publishers.

Ferrell, W. R. (1966). Delayed force feedback. Human Factors, October, 449-455.

Flach, J. M. (1994). Situational awareness: The emperor's new clothes. In M. Mouloua and R Parasuraman (Eds.), Human Performance in Automated Systems: Current Research and Trends. Hillsdale NJ: LEA.

Friedman, J., Globus, G., Huntley, A., Mullaney, D., Naitoh, P., and Johnson, L. (1977). Performance and mood during and after gradual sleep reduction. Psychophysiology, 14, 245-250.

Frese, M. & Okenek K. (1984) Reasons to leave shiftwork and psychosocial and psychosomatic complaints of former shiftworkers. Journal of Applied Psychology, 69, 509-541.

Folkard, S. (1975) Diurnal variation in logical reasoning. British Journal of Psychology, 66, 1-8.

Folkard, S. Knauth, P., Monk, T. H. & Rutenfranz, J. (1976) The effect of memory load on the circadian variation in performance efficiency under a rapidly rotating shift system. Ergonomics, 19, 479-488.

Folkard, S., Wever, R. A. & Wildgruber, C. M. (1983) Multi-oscillatory control of circadian rhythms. Nature, 305, 223-226.

Foard, G., Kemler, A. T., & Nelson, L (1989). Human Performance Modeling in Display Design, (Tech. note 3-88). Aberdeen Proving Ground, MD: Human Engineering

Laboratory

Garner, W. R. (1970). The stimulus in information processing. American Psychologist, 25, 350-358.

Garner, W. R. (1974). The Processing of Information and Structure, New Jersey: Lawrence Erlbaum.

Garner, W. R. (1981). The analysis of unanalyzed perceptions. In M. Kubovy & J. R. Pomerantz (Eds.) Perceptual Organization, New Jersey: Lawrence Erlbaum.

Gentner, D., & Stevens, A. L. (1983). Mental Models. New Jersey: Lawrence Erlbaum.

Gibson, J. J. (1979). The Ecological Approach To Visual Perception. Boston: Houghton Mifflin.

Gillis, J. S. (1969). Schizophrenic thinking in a probabilistic situation. Psychological Record, 19, 211-224.

Gilson, R. D., Garland, D. J., & Koonce, J. M. (1994). Situational Awareness in Complex Systems. Daytona Beach: Embry-Riddle Aeronautical University Press.

Gordon, I. (1990). Theories of Visual Perception. New York: John Wiley.

Hamilton, P., Wilkinson, R. T., and Edwards, R. S. (1972). A study of four days partial sleep deprivation. In: W. P. Colquhoun (Ed.) Aspects of Human Efficiency. London: English University Press, 101-113.

Hammond, K. R., Hamm, R. M., Grassia, J. & Pearson, T. (1987). Direct comparison of the efficacy of intuitive and analytical cognition in expert judgments. IEEE Transactions on Systems, Man and Cybernetics, SMC-17, 5, 753-770.

Hammond, K. R., & Adelman, L. (1976). Science, values, and human judgment. Science, 194, 389-396.

Hammond, K. R., McClelland, G. H., & Mumpower, J. (1980). Human Judgment and Decision Making: Theories, Methods and Procedures. New York: Praeger.

Hammond, K. R., & Wascoe, N. E. (1980). Realizations of Brunswik's Representative Design. San Francisco: Jossey-Bass Inc., Publishers

Hammond, K. R. (1966). The Psychology of Egon Brunswik. New York: Holt, Rinehart, and Winston.

Hammond, K. R., Hamm, R. M. & Grassia, J. (1986). Generalizing over conditions by

combining the multitrait-multimethod matrix and the representative design of experiments. Psychological Bulletin, 100, 257-269.

Hammond, K. R. (1980). The Integration of Research in Judgment and Policy. Center For Research on Judgment and Policy, Report No. 226.

Hancock, P. A. (1984). Environmental stressors. In J. S. Warm (Ed.) Sustained Attention in Human Performance, New York: John Wiley.

Hedges, J. N. (1971) A look at the 4-day work week. Monthly Labor Review, 33-37.

Heslegrave, R. J., and Angus, R. G. (1982). Sleep Loss and Continuous Work. NATO Report, Defense and Civil Institute of Environmental Medicine.

Heslegrave, R. J., and Angus, R. G. (1985). The effects of task duration and work session location on performance degradation induced by sleep loss and sustained cognitive work. Behavioral Research Methods, Instruments and Computers, 17, 592-603.

Hoddes, E., Zarcone, V., Smythe, H., Phillips, R., and Dement, W. C. (1973). Quantification of sleepiness: a new approach. Psychophysiology, 10, 431-436.

Hollenbeck, J. R, Ilgen, D. R., Sego, D. J., Hedlund, J. Major, D. A., & Phillips, J. (1995). Multilevel theory of team decision making: Decision making in teams incorporating distributed expertise. Journal of Applied Psychology, 80, 292-316.

Johnson-Laird, P.N. (1990). Human thinking and mental models. In: K.A. Mohyeldin Said, W.H. Newton-Smith, R. Viale, and K.V. Wiles (Eds.) Modelling the Mind. Oxford: Clarendon Press, 155-170.

Johnson-Laird, P.N. (1975). Models of deduction. In: R. Falmagne (Ed.) Reasoning: Representation and Process. New Jersey: Lawrence Earlbaum.

Johnson, L. C. (1982). Sleep deprivation and performance. In: W. B. Webb (Ed.) Biological Rhythms Sleep and Performance, New York: John Wiley, 111-138.

Johnson, L. C., and Naitoh, P. (1974). The Operational Consequences of Sleep Deprivation and Sleep Deficit. AGARDograph AF-193, NATO Advisory Group for Aerospace Research and Development, London, England.

Kahneman, D., Slovic, P., & Tversky, A. (1982). Judgment Under Uncertainty:

Heuristics and Biases. Cambridge: Cambridge University Press.

Kelly, R. J. & Schneider, M. F. (1982) The twelve-hour shift revisited: Recent trends in the electric power industry. Journal of Human Ergology, 11, Suppl.: 369-384.

Kleitman, N. (1963) Sleep and Wakefulness. Chicago: University of Chicago Press.

Kleinmuntz, D. N. (1985). Cognitive heuristics and feedback in a dynamic decision environment. Management Science, 31, 680-702.

Krueger, G. P. (1989). Sustained work, fatigue, sleep loss and performance. Work and Stress, 3, 129-142.

Larkin, J. H. (1979). Information processing and science instruction. In J. Lockhead & J. Clement (Eds.), Cognitive Process Instruction. Philadelphia: Franklin Institute Press.

Larkin, J. H, McDermott, J., Simon, D. P. & Simon, H. D. (1980). Expert and novice performance in solving physics problems. Science, 208, 1335-1342.

Mahan, R. P. (1994) Stress induced strategy changes toward intuitive cognition: A cognitive continuum framework approach. Human Performance, Vol 7, 3.

Mahan, R. P. (1991a). Time-on-task effects within a probabilistic task environment. Perceptual and Motor Skills, 72, 307-314.

Mahan, R. P., Carvalhais, A., & Queen S. (1990). Sleep reduction in night-shift workers: Is it sleep deprivation or a sleep disturbance disorder? Perceptual and Motor Skills, 70, 723-730.

Mahan, R. P. (1991b). The effects of extended work on multi-dimensional judgments of a manufacturing process. International Journal of Human Factors in Manufacturing, 1, 155-165.

Mahan, R. P. (1992). Effects of task uncertainty and continuous performance on knowledge execution in complex decision making. International Journal Computer Integrated Manufacturing, 5, 58-67.

Newell, A. (1973). Production systems: models of control structures. In W. G. Chase (Ed.), Visual Information Processing (pp. 283-308). New York: Academic Press.

Newell, A. & Simon, H. A. (1972). Human Problem Solving. New Jersey: Prentice-Hall.

Payne, J. (1982). Contingent decision behavior. Psychological Bulletin, 92, 382-402.

Pomerantz, D. (1981). Object Perception, In M. Kubovy & J. R. Pomerantz (Eds.) Perceptual Organization, New Jersey: Lawrence Erlbaum.

Rasmussen, J. (1981). Models of mental strategies in process control. In J. Rasmussen and W. Rouse (Eds.), Human Detection and Diagnosis of System Failures. New York: Plenum Press.

Rasmussen, J. (1986). Information Processing and Human Machine Interaction. New York: North Holland.

Rasmussen, J. (1988). A cognitive engineering approach to the modeling of decision making and its organization in process control, emergency management, CAD/CAM, office systems, and library systems. In: W. B. Rouse (Ed.), Advances in Man-Machine Systems Research. 4, Greenwich, CT.: JAI Press.

Rouse, W. B. (1977). A theory of decision making in stochastic estimation tasks. IEEE Transactions on Systems, Man, and Cybernetics, SMC-7, 274-283.

Rouse, W. B. (1980). Systems Engineering Models of Human-Machine Interaction. New York: North Holland.

Rosa, R. R., Wheeler, D. D., Warm, J. S., and Colligan, M. J. (1985). Extended workdays: effects of performance and ratings of fatigue and alertness. Behavioral Research Methods, Instruments and Computers, 17, 6-15.

Sanderson, P. M., Flach, J. M., Buttigieg, M. A. & Casey, E. J. (1989). Object displays do not always support better integration task performance. Human Factors, 31, 183-198.

Sheridan, T. B. (1976). Toward a general model of supervisory control. In: T.B. Sheridan and G. Johannsen (Eds.) Monitoring Behavior and Supervisory Control. New York: Plenum Press, 271-282.

Sheridan, T. B. (1987). Supervisory control. In: G. Salvendy (Ed.) Handbook of Human Factors, New York: John Wiley, 971-989.

Shiffrin, R. M. & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. Psychological Review, 84, 127-190.

Simon, H. A. (1978). Rationality as a process and product of thought. American Economic Review, 68, 1-16.

Simon, H. A. (1976). Administrative Behavior: A Study of Decision Making Processes in Administrative Organizations, (3rd edition), New York: Free Press.

Simon, H. A. (1957). Models of Man: Social and Rational. New York: John Wiley.

Slatter, P. E. (1987). Building Expert Systems: Cognitive Emulation. Chichester: Ellis Horwood.

Slovic, P., Fischhoff, B. & Lichtenstein, S. (1977). Behavioral decision theory. Annual Review of Psychology, 28, 1-39.

Tepas, D. I., and Mahan, R. P. (1989). The many meanings of sleep. Work and Stress, 3, 93-102.

Tucker, L R. (1964). A suggested alternative formulation in the development of Hursch, Hammond and Hursch and by Hammond, Hursch and Todd. Psychological Review, 71, 528-530.

Tversky, A. (1973). Elimination by aspects. Psychological Review, 79, 281-299.

Tversky, A., and Kahneman, D. (1974). Judgment under uncertainty: Heuristics and Biases. Science, 185, 1124-1131.

U.S. Congress, Office of Technology Assessment, (1991). Biological Rhythms: Implications for the Worker. OTA-BA-463, Washington, D.C.: Government Printing Office.

von Winterfeldt, D. & Edwards, W. (1986). Decision Analysis in Behavioral Research. Cambridge: Cambridge University Press.

Warm, J. S. (Ed.) (1984). Sustained Attention in Human Performance. New York: John Wiley, 1-14.

Wickens, C. D. (1984). Engineering Psychology and Human Performance. Glenview, Illinois: Scott, Foresman and Company.

Wickens, C. D., & Andre, A. D. (1990). Proximity compatibility and information display: effects of color, space, and objectness on information integration. Human Factors, 32, 61-78.

Wright, J. C. & Murphey, G. L. (1984). The utility of theories in intuitive statistics: the

robustness of theory-based judgments. Journal of Experimental Psychology: General, 113, 301-322.

Wilkinson, R. T. (1961). Interaction of lack of sleep with knowledge of results, repeated testing and individual differences. Journal of Experimental Psychology, 62, 263-271.

Wilkinson, R. T. (1969). Sleep deprivation: performance tests for partial and selected sleep deprivation. Progress in Clinical Psychology, 8, 28-43.

Williams, H. L., Lubin, A., and Goodnow, J. J. (1959). Impaired performance with acute sleep loss. Psychological Monographs, 73, (14, Whole No. 484).

Williams, H. L., and Lubin, A. (1967). Speeded addition and sleep loss. Journal of Experimental Psychology, 73, 313-317.

Woods, D. (1988). Coping with complexity: The psychology of human behavior in complex systems. In L. P. Goodstein, H. B. Anderson, and S. E. Olsen (Eds.), Tasks, Errors and Mental Models. Philadelphia, PA: Taylor and Francis.

AN INVESTIGATION OF THE PRACTICALITY OF THE RAPID ITS DEVELOPMENT
SHELL (RIDES) USE IN THE COLLEGIATE ENVIRONMENT

Patrick R. Mattson
Assistant Professor
Department of Technology - Aviation

St. Cloud State University
720 Fourth Ave South
St. Cloud, MN 56301-4498

Final Report for:
Summer Faculty Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Armstrong Laboratory
Brooks AFB, TX

September 1995

AN INVESTIGATION OF THE PRACTICALITY OF THE RAPID ITS DEVELOPMENT
SHELL (RIDES) USE IN THE COLLEGIATE ENVIRONMENT

Patrick R. Mattson
Assistant Professor
Department of Technology - Aviation
St. Cloud State University

Abstract

The practical use of interactive computerized instructional authoring tools by post secondary teachers holds exciting promise for the enhancement of the educational process. However, before this is a reality, a suitable, powerful authoring shell must be made available to allow the instructor to work without a computer programmer. The Rapid Intelligent Tutoring System (ITS) Development Shell (RIDES) (Munro, 1995) holds such promise of being that tool. RIDES, a product of R&D, is a new computer-based training (CBT) authoring shell tool; at this point it is a prototype tool for the production and delivery of learner interactive training through a simulation-based environment on a 80486 or higher class computer running under the UNIX operating system. This research investigated the RIDES authoring shell as to its value, ease of use and the time and effort required in constructing a moderately complex, working simulation tutor by a Subject Matter Expert (SME) in the Air Traffic Control Training (ATCT) domain. Future study should be conducted comparing several commercial CBT authoring packages with RIDES to test the pedagogical effectiveness of the ATCT Basic Tutor, gather baseline data against traditional instruction in an actual academic setting, and validate the effectiveness of RIDES by an SME, i.e. the college teacher and its related benefits.

TABLE OF CONTENTS

Introduction	4
Discussion	5
Procedures	9
Conclusions and Recommendations	11
References	14

INTRODUCTION

RIDES is currently under development by the University of Southern California, Behavioral Technology Laboratories and the USAF's Armstrong Laboratory. Can the RIDES authoring shell tool be used by the SME (as discussed in this paper, a domain expert with minimal or no assistance from a computer programmer) to develop a robust application in minimum time using available computer hardware that would be useful in a normal collegiate academic setting? An Intelligent Tutoring System (ITS) capable of helping students learn complex problem-solving tasks requires extremely flexible human-machine interfaces (Burns & Parlett, 1991). Is RIDES the answer? Does the software allow users with *some* computer knowledge and expertise to build a computer tutor that would be useful in a normal classroom setting? Could this tutor then be used to supplement instruction, possibly allowing the instructor to advance his/her students farther along the learning process than if they were using traditional teaching methods (group lecture, discussion, paper/pencil testing, etc.)? This research and paper attempts to answer these questions.

According to Instructional Designs (1995), research has shown that computer-based interactive education and instruction:

- increases the consistency of learning
- increases trainee [student] retention of materials by 25% to 50%
- is more cost effective to deliver than traditional classroom training
- is available day or night
- reduces delivery variance
- reduces training time by 20% to 40%

We believe that Artificial Intelligence (AI) based tutor systems have great potential for training applications for ATC (Means, Mumaw, Roth, Schlager, McWilliams, Gagne', Rice, Rosenthal, & Heon, 1988). The air traffic control training (ATCT) domain is ideal for interactive computer-based instruction due to the vast amount of information the student must quickly absorb and the requirement for the student to visualize large amounts of spatially related data.

This paper provides anecdotal results of the author's experience designing, building and initial testing of a RIDES ATCT Basic Tutor.

DISCUSSION

The goal of authoring systems is to give relative computer novices a software toolkit to take advantage of the power of computers for designing instruction (Shute & Psotka, 1994). RIDES provides a powerful computer platform which an SME can use to develop robust teaching tools for the classroom and/or lab. Kurt Cagle (1995) writing in the August 1995 issue of Computer Shopper, discusses how to size up an authoring tool in this way:

Actually, the reality of multimedia [CBT] creation is considerably more complex. If you try to create a multimedia application without the proper tools, it'll be you, not the project, that ends up in the pressure cooker. At any level, an authoring tool's *raison d'être* is to integrate various computer media.

An authoring package should also include some form of scripting language, if only an iconic one. Finally, the product you create with an authoring tool must be transportable without having to include the authoring tool.

The oldest class of multimedia authoring tools, and in many ways, still the most powerful, are the scripted authoring tools that evolved from programming languages. Technically, a scripting language is a programming language; it follows the same algorithmic structures as a development language such as C or Pascal. The difference between the two involves what the language is used for and how tightly integrated it is to the underlying operating system.

Because the power of these authoring environments resides in their scripting languages, they tend to be the least user-friendly, taking longer to master than property-based or iconic tools.

Generally, programs for development of CBT provide a graphical interface to allow the author to concentrate on designing useful applications for the student. The transition from novice to experienced developer, a process that an SME may experience several times, depends on how intuitive the program appears to that person. Although Cagle indicates that scripting is the more powerful way to develop CBT applications, programming detracts from the main goal of designing the application. Several companies offer CBT tools which allow the SME to develop applications that incorporate problem solving and coaching but lack true deep-level simulation capabilities. Computer-Based Instruction (CBI) [and CBT] does have a tremendous potential to help the world meet future educational challenges (Schlechter, 1991). Expectations of the SME as to how well an authoring shell mimics his or her past experience with computer

systems must be included in the design of an effective tool. In turn, the productivity of a developer could be enhanced by the right product to supplement classroom instruction and appears to be dependent on the extent to which they view the ease of use and capability of the authoring shell.

Air traffic control (ATC) employment with the U. S. Air Force (USAF) and the Federal Aviation Administration (FAA) is a very demanding but rewarding task requiring well trained persons. Current employment and recruiting practices dictate that the applicant be hired based on the potential or aptitude for completing a training program which has remained virtually unchanged since the 1960's. One exception to this is the inclusion of high-fidelity simulation equipment (designed primarily for students who have been taught the basic concepts) into the training lab. Very little investigation or development work has been done in the area of initial air traffic control training tutor/simulation equipment. During initial training the student is taught the basics of air traffic control; the student is then given the opportunity to demonstrate his/her ability to apply this knowledge in a simulated ATC environment over a prolonged period of time. Using the Instructional Systems Development (ISD) method of teaching, both the USAF and FAA attempt to ensure that controller trainees have a complete grasp of air traffic control procedures for their assigned option.

Bonar (1990) indicated that three kinds of domain expertise should be considered: Concepts, representations, and rules for referring to situations in the world; the interface needs to embody all three aspects - that is, provide the integration that is natural in an expert's performance.

Figure 1 shows how RIDES does include Bonar's definition.

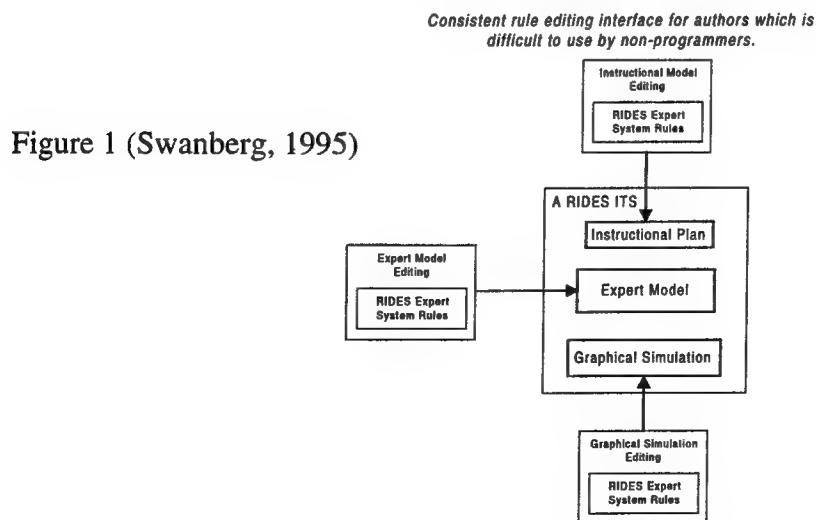


Figure 1 (Swanberg, 1995)

A central issue in designing intelligent tutoring systems is the nature of the knowledge of the target domain (Reiser, Kimberg, Lovett, & Ranney, 1992). Over the years, extensive evidence has accumulated to show that the form of mental picture the controller has of the traffic is influenced by the kinds of training that the controller originally received and by the forms of information presentation and of computer assistance that were available during that training (Hopkin, 1988). These requirements make the ATCT domain ideal for interactive computer-based simulation (CBT).

Current ATC training practices force students to memorize "parts" and then the instructor (one-to-one) spends months to years reiterating the same information, trying to assist the trainee/student to see how these fit together. The RIDES ATCT tutor, after using interactive knowledge drills, helps the student see all the "parts" together and it allows for repeated examination to ensure the student learns the "parts", and, by teaching these in the context of a simulated environment, it also gives them a feel for real world type action. This approach to teaching early phases of ATC concepts is new and worthy of future development. Figure 2 details the lesson planning for the tutor and represents a new way of learning critical knowledge essential to a controller's success on the job. The objective is to enhance the learning of basic ATC knowledge for faster acquisition and better retention. This process will allow the student to see relationships between distinct but interrelated pieces of information vital to this domain i.e., map knowledge, flight progress strips, and separation rules. Students will be able to internalize the information thereby allowing them to acquire their own controller technique faster.

Air Traffic Control Training - RIDES Tutor

Lesson Flow

08/03/95

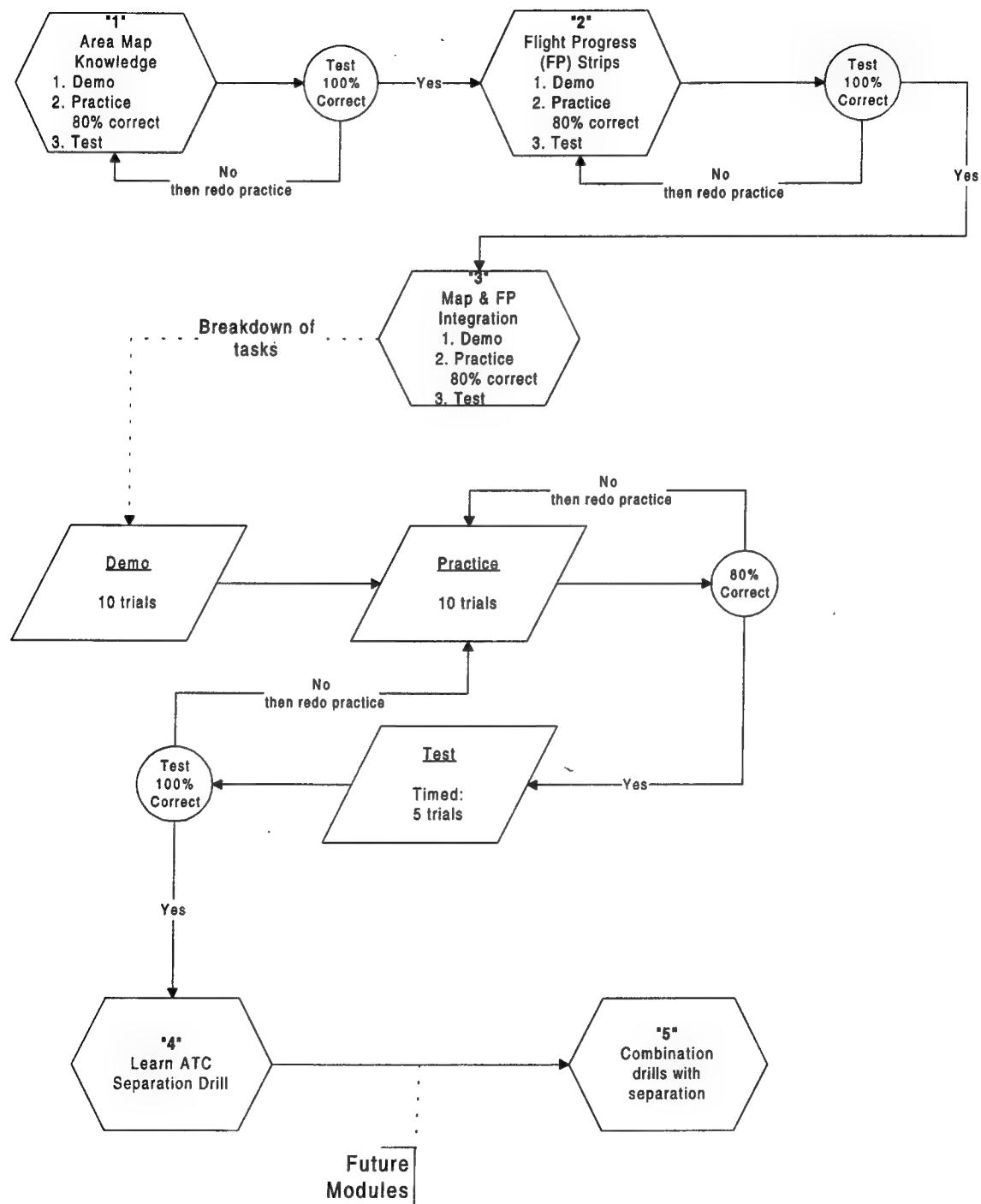


Figure 2

PROCEDURES

Over a period of nine weeks, this SME built an ATCT tutor using RIDES. It presented the area map and route-of-flight information together (Figure 3 shows a multi-scene view). Development was slow due to the initial task of learning a new computer operating system and user interface. Fleming & Stueck, (1990) proposed that three components or models be present in an effective tutor: Expert, Student, and Instructional. This RIDES tutor embodies all of these models.

The tutor was designed to closely match what a human ATC trainer would require of the trainee before allowing him/her to proceed on to more complex tasks. The tutor teaches the student using an interactive two-dimensional rather than a one dimensional (paper and pencil) tool to learn all phases of the initial instruction. The aircraft object is driven by an animation script; to show an airplane "flying" on the correct route, the student is asked to click on the correct points making up that route from the flight progress strip information. This is the same type of information that the controller would work with in real life. This enables the student to acquire the "mental model" (see Rowe, Cooke, Hall, & Halgren, 1995--in press--for a definition of mental models) of how aircraft information and the area control map fit together. During the earliest phases of instruction, the student is given the chance to "fly" the aircraft through their assigned airspace. This helps students see the impact of their control instructions and should allow a faster grasp of the more complex concepts of aircraft performance, separation rules, and control procedures and methods. Each module can act as a stand alone lesson or become part of a fully interactive air traffic control tutor.

The tutor also provides for a decrease in instructor bias as the student learns and practices "textbook" rules and methods. Several actual air traffic control training observations leads this author to conclude that students who develop their own technique early and are allowed to practice with little instructor influence get a "mental model" faster than those who practice technique while being heavily influenced by a human instructor. The approach used in the ATCT tutor allows the student to internalize a "mental model" of three essential aspects of air traffic control - flight progress strips, how the area map relates to the aircraft data, and how the aircraft flies the route.

One view (scene) of the ATCT RIDES Tutor

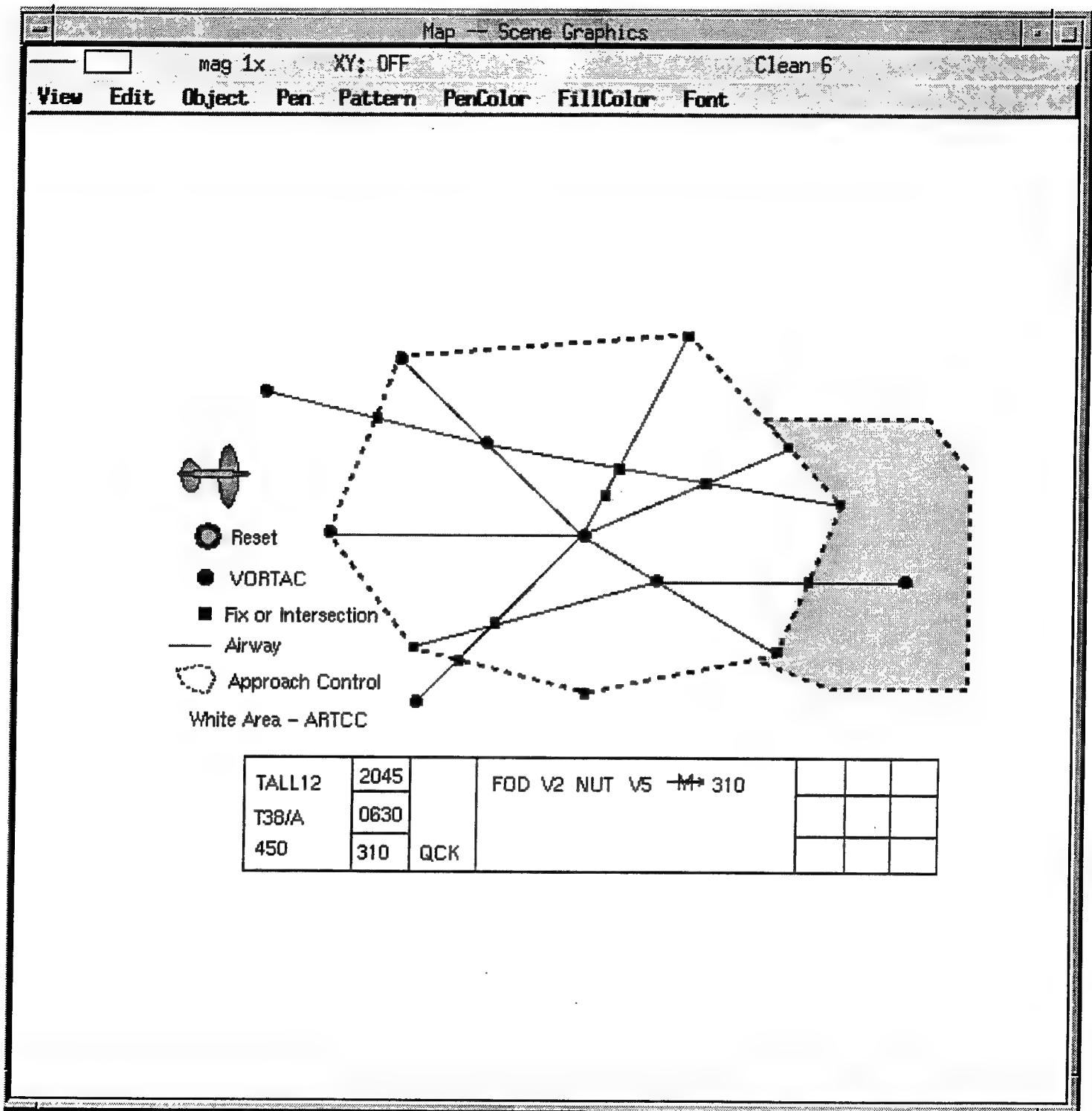


Figure 3

CONCLUSIONS

Building an acceptable ITS for this domain is another story. Achieving the measures outlined by Instructional Designs would greatly enhance controller productivity and shorten the time spent in initial training but RIDES in the hands of the non-programming SME is not the answer at this time. The following is offered as a means for the reader to quickly grasp the author's overall evaluation of the RIDES software program. These observations are based on my ten plus years of working with computers and intimate familiarity with an MS-DOS, Windows based computer operating system and limited exposure to Apple's Macintosh system. During the first three weeks, most of the development time was spent trying to learn an unfamiliar operating system.

Drawbacks to the program's useability include:

- no libraries or templates of objects or instruction
- too much scripting of program code needed
- relation between Author and Student modes is not automatic
- on-line help (ROSS) tutor is not built into RIDES as a help function
- most instructors are not inclined to do any programming
- no help menus
- no mini-manual of steps to help build application; manual is too complicated
- requires knowledge of software logic & flow
- UNIX is normally viewed by post secondary teachers as a mainframe environment
- X-windows graphical user interface is not intuitive

Good to excellent aspects of this program include:

- automated instruction generator (patterned exercises)
- authoring tool supports deep-level simulation
- completed tutor supplements instructor in information-rich domains
- objects can be "brought to life" using animation and scripting techniques
- windowing placement of the information to achieve an efficient format
- allows for true deep-level simulation
- CBI is proven to enhance the instructional process
- graphical objects are linked to instructional logic
- powerful 32-bit UNIX based environment
- smart graphical objects

The drawbacks need to be corrected before the program can gain acceptance in the post secondary community. One possible solution would be to port the program code to a more widely accepted, desktop operating system such as Microsoft Windows NT or 95.

According to a survey of their readers, Computer Shopper (June, 1995) magazine (raw data unavailable) reported the following usage:

<u>Percent of users</u>	<u>Desktop Operating System</u>
81.0	DOS with Windows 3.1
8.2	IBM OS/2
3.9	DOS alone
3.1	Apple Macintosh
2.0	Windows NT
1.8	UNIX/Other

The same article indicated that 85.2% were planning to switch to Windows 95 soon after it was released. It is this author's opinion that a UNIX-based RIDES authoring shell will not gain wide acceptance as the program is too frustrating to develop even the most rudimentary application.

If scripting (Cagle, 1995) is the more powerful way to build a tutor then RIDES is an excellent shell to use but the average SME will have much difficulty concentrating on design if they are not skilled in basic programming techniques. Cognitive directness (in design) involves minimizing mental transformations that a user must make; even small cognitive transformations by a user take away effort from the intended task (Hix & Harston, 1993) RIDES is awkward and not natural to use; it does not follow "normal" computer learned behavior, such as Windows mouse movements. While some parts of the program (patterned exercises utility which allowed a few mouse clicks and the "hard work" was done by the computer) are excellent, the impression is that RIDES was developed for the programmer (giving its powerful scripting language) and not for the SME who has very little computer expertise. I would not recommend that an SME with less than expert computer knowledge level attempt to use this software. RIDES Unix-based software is powerful, but it is not user friendly; porting the software to a different, more widely accepted, desktop operating system environment (Windows NT, 95 or OS/2) would greatly enhance the viability of this product. However, in order for RIDES to be accepted by post secondary faculty in their classrooms, the program must evolve as shown in Figure 4 to the point that a programmer is needed only to assist with technical support to help with the most complex details of the software.

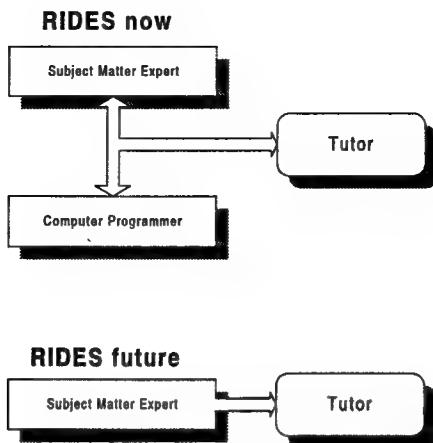


Figure 4

Even though RIDES is difficult to use and develop a realistic application for use in the classroom by the non-computer expert, future research is proposed in the following areas: complete modules 4 & 5 of the ATCT program in UNIX, compare commercial off-the-shelf authoring shells (MS Windows based), gather other SME input as to their impressions of RIDES, gather baseline data in an actual classroom setting, and test and validate the pedagogical effectiveness of RIDES by this SME.

REFERENCES

Bonar, J. (1990). *Interface Architectures for Intelligent Tutoring Systems*. In H. Burns, J. Parlett, & C. Luckhart, C. (Eds). Proceedings of the Second Intelligent Tutoring Systems Research Forum (AFHRL-TP-89-31). (pp39-53) Brooks AFB, TX: Training Systems Division, Air Force Human Resources Laboratory.

Burns, H. L., & Parlett, J. (1991). *The Evolution of Intelligent Tutoring Systems: Dimensions of Design*. In Burns, H. L., & Parlett, J., & Luckhart-Redfield, C. (Eds). Intelligent Tutoring Systems: Evolution in Design. (pp1-11) Hillsdale, NJ: Lawrence Erlbaum Associates.

Cagle, K. (August, 1995). *Multimedia Authoring Tools for Windows* (On-Line) Available http://www.zdnet.com/~cshopper/features/9508/sw1_9508/index.html Ziff-Davis Publishing Company. New York, NY.

Clark, R. C. (1989). *Developing Technical Training: A Structured Approach for the Development of Classroom and Computer-Based Instructional Materials*. Reading, MA: Addison-Wesley Publishing Company, Inc.

Clark, R. C. (1995). *21st Century Human Performance*. Training. Vol 32, No. 6, 85-90.

Computer Shopper. (June, 1995). *On-Line Survey Results*. Ziff-Davis Publishing Company. New York, NY. Vol 15, No. 6, Issue 183, 47.

Deaton, W. L. (1991). *CBT and Higher Education*. In T. M. Shlechter (Ed.) (1991). Problems and Promises in Computer-Based Training. (pp. 215-231) Norwood, NJ: Ablex Publishing Corporation.

Fleming, J. L. & Stueck, K. (1990). *Intelligent Tutoring Systems: A Taxonomy of Evaluation Issues* (AFHRL-TP-89-79). Brooks AFB, TX: Training Systems Division, Air Force Human Resources Laboratory.

Grabowski, B. & Vroms, K. (1994). *ID and the Use of Authoring Systems*. In W. D. Milheim (Ed.). Authoring-Systems Software for Computer-Based Training. (pp. 3-22) Englewood Cliffs, NJ: Educational Technology Publications.

Hix, D. & Harston, H. R. (1993) *Developing User Interfaces: Ensuring Usability Through Product & Process*. New York: John Wiley & Sons, Inc.

Hopkin, V. *Training Implications of Technological Advances in Air Traffic Control*. (1988) In The Proceedings of the Symposium on Air Traffic Control Training for Tomorrow's Technology (pp. 6-26). Sponsored by the Federal Aviation Administration, December 6 & 7, 1988. Oklahoma City, OK.

Instructional Designs, Inc (1995) *Instructional Designs* (On-Line) Available
<http://www.SafetyOnline.net/training/instdsgn/instdsgn.htm>. Greer, SC: Craftsman Court

Means, B., Mumaw, R., Roth, C., Schlager, M., McWilliams, E., Gagne', E., Rice, V., Rosenthal, D., & Heon, S. (1988) *ATC Training Analysis Study: Design of the Next-Generation ATC Training System* (Technical Report, OPM Work Order 342-036). Washington, DC, Federal Aviation Administration.

Munro, A. (1993). *An Introduction to RIDES Authoring*. Unpublished manuscript, University of Southern California, Redondo Beach, CA: Behavioral Technology Laboratories.

Munro, A. (1994). *RIDES Authoring Reference*. Unpublished manuscript, University of Southern California, Redondo Beach, CA: Behavioral Technology Laboratories.

Munro, A. (1995). *RIDES Authoring Software version 3.5.4*. [Computer Software]. University of Southern California, Redondo Beach, CA: Behavioral Technology Laboratories.

Reiser, B., Kimberg, D., Lovett M., & Ranney, M. (1992) *Knowledge Representation and Explanation in GIL, An Intelligent Tutor for Programming*. In J. Larkin, J. H., & Chabay, R. W. (Eds.). Computer-Assisted Instruction and Intelligent Tutoring Systems: Shared Goals and Complementary Approaches. (pp 111-149). Hillsdale, NJ: Lawrence Erlbaum Associates.

Richardson, J. J. (1983). *Artificial Intelligence: An Analysis of Potential Applications to Training, Performance Measurement, and Job Performance Aiding* (AFHRL-TP-83-28). Lowery AFB, CO: Training Systems Division, Air Force Human Resources Laboratory.

Rowe, A. L., Cooke, N. J., Hall, E. P., & Halgren, T. L. (in press) *Toward an On-Line Knowledge Assessment Methodology: Building on the Relationship Between Knowing and Doing*.

Shlechter, T. M. (1991). *Promises, Promises, Promises: History and Foundations of Computer-Based Training*. In T. M. Shlechter (Ed.) (1991). Problems and Promises in Computer-Based Training. (pp. 1-20) Norwood, NJ: Ablex Publishing Corporation.

Shute, V. J. & Psotka, J. (1994). *Intelligent Tutoring Systems: Past, Present, and Future* (AFHRL-TP-1994-0005). Brooks AFB, TX: Human Resources Directorate, Manpower and Personnel Research Division, Human Resources Laboratory.

Spector, J. M. (1990). *Designing and Developing an Advanced Instructional Design Advisor* (AFHRL-TP-90-52). Brooks AFB, TX: Training Systems Division, Air Force Human Resources Laboratory.

Stepenson, S. D. (1990). *Role of the Instructor in Maximizing Academic Achievement in Computer-Based Training* (AFHRL-TP-90-24). Brooks AFB, TX: Training Systems Division, Air Force Human Resources Laboratory.

Stoker, P., Hunter, D. R., Batchelor, C. L., & Curran, L. T (1987). *Air Traffic Controller Trainee Selection* (AFHRL-TP-87-19). Brooks AFB, TX: Human Resources Directorate, Manpower and Personnel Research Division, Human Resources Laboratory.

Swanberg, C. (1995). *Simulators in Intelligent Environments: RIDES*. (Unpublished Document). Brooks AFB, TX: Training Systems Division, Air Force Human Resources Laboratory.

Towne, D. M. & Munro, A. (1991). *RAPIDS: A Simulation-Based Instructional Authoring System for Technical Training* (AFHRL-TR-90-49). Brooks AFB, TX: Training Systems Division, Air Force Human Resources Laboratory.

**Repetitive Sequence Based PCR (REP-PCR): An Epidemiological Study of a
Streptococcus Pneumoniae Outbreak in Pennsylvania**

**Sandra L. McAlister
Associate Professor
Department of Biology**

**Stonehill College
North Easton, MA 02357**

**Final Report for:
Summer Faculty Research Program
Armstrong Laboratory**

**Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base
Washington, D.C.**

and

Armstrong Laboratory

October 1995

**REPETITIVE SEQUENCE BASED PCR (REP-PCR): AN EPIDEMIOLOGICAL STUDY OF
A *STREPTOCOCCUS PNEUMONIAE* OUTBREAK IN PENNSYLVANIA**

Sandra L. McAlister
Associate Professor
Department of Biology
Stonehill College
North Easton, Ma. 02357

ABSTRACT

Rep-PCR is used as an epidemiological tool in this blind study of a potential outbreak of *Streptococcus pneumoniae* from Pennsylvania. BOXA1R primer was attached to repetitive, interspersed, non-coding elements of pneumococcal DNA. Areas of DNA between the successive primers were amplified by PCR, and amplicon was separated into distinct bands of DNA by agarose gel electrophoresis. Each sample of pneumococcus was fingerprinted and compared, both by visual and computerized methods, to yield a strain specific identity for each class of organism provided. The BOXA1R Rep-PCR method herein is proposed as a rapid, sensitive, cost-effective procedure ideal for use in epidemiological studies of suspected pneumococcal outbreaks.

ACKNOWLEDGEMENTS

I wish to thank the Air Force Office of Scientific Research for supporting this program and for the opportunity to conduct this study. I am particularly indebted to Ferne K. McCleskey for the use of her laboratory space and equipment, her unending encouragement and her sound advise and tutelage. Without the assistance of Dr. Vito DelVecchio and two of his graduate students, Beth Ann McBrearty and Catherine Healy, I most certainly would not have made the gains in technique and progress toward optimizing this new procedure to fingerprint the DNA of a medically important bacterial pathogen. The time I have spent here and the successes (and setbacks) experienced have been invaluable.

**REPETITIVE SEQUENCE BASED PCR (REP-PCR): AN EPIDEMIOLOGICAL STUDY OF
A *STREPTOCOCCUS PNEUMONIAE* OUTBREAK IN PENNSYLVANIA**

Sandra L. McAlister

INTRODUCTION

Streptococcus pneumoniae (*S. pneumoniae*) is a gram-positive, catalase-negative, alpha-hemolytic, lanceolate-shaped organism which frequently appears as a diplococcus. The organism in its pathological role causes pneumonia, meningitis and otitis media in children, the elderly and those who are immunocompromised (4, 10). The recent escalation of penicillin-resistant strains of pneumococcus in the United States seems to foreshadow the dramatic spread of resistant organism seen in Europe, Australia and South Africa over the past twenty years (10, 12) .

The epidemiology of both the penicillin susceptible and penicillin resistant types of the pneumococcus has been studied by many methods from conventional serotyping to various of the molecular strategies available in the past decade (2, 5, 6, 8, 13, 14). The goal of these strategies is to find a rapid, repeatable, and reliable method of identifying and distinguishing between penicillin-susceptible and penicillin-resistant strains. With this type of information at the disposal of the physician and epidemiologist, rapid knowledge of and timely curtailment of potential outbreaks, identification of "carriers" and proper antibiotic treatment for the infected is possible in a timeframe never before possible. The 'front lines' of recognition of a possible outbreak is the clinical laboratory. Samples arriving from many sites and sources, will be rapidly assessed by Rep-PCR and identified specifically as resistant or non-resistant pneumococci, and will be reported to the clinician within days of the receipt of the first sample.

DISCUSSION OF PROBLEM:**NATURAL HABITAT:**

S. pneumoniae display both a commensal and a parasitic relationship with humans. It is a frequent colonizer of skin and mucous membranes. It is also frequently found as a co-habitant of the "normal flora" of the respiratory tract and other specific tissues. While not to be considered "normal" flora, *S. pneumoniae* colonizes approximately 35% of the population without the appearance of clinical symptoms (10).

More commonly, *S. pneumoniae* is a virulent pathogen, exclusive to humans, which produces tissue damage to lung, middle ear tissues, the meninges of the central nervous system, and other discrete tissues. It is spread from person-to-person by an aerosol route, adopting the naso- pharyngeal mucosa (10).

CONVENTIONAL DIAGNOSTIC TESTING:

Among the clinical laboratory tests used to identify *S. pneumoniae* are typical culture and sensitivity and gram-staining. After 24 hours on blood or chocolate agar, the organism will show ample growth and alpha-hemolysis. Conducting optochin, bile solubility and Pneumoslide (BBL) tests will differentiate between *S. pneumoniae* and *S. viridans* , providing a confirmation of genus and species. To determine its antibiotic sensitivities requires an additional twenty-four hours. To determine the capsular serotyping, a time-consuming and labor intensive process, is usually beyond the capabilities of the routine microbiology laboratory. It may take months to perform a comprehensive identification of the pneumococcal samples by conventional diagnostic laboratory methods.

ANTIBIOTIC RESISTANCE:

S. pneumoniae strains fall into three broad categories with regard to antibiotic resistance: penicillin-susceptible, penicillin-resistant, and relatively penicillin-resistant strains. Of the three the most clinically perilous is the relatively resistant strains. These are defined as those which are somewhat susceptible to penicillin in vitro, but, when administered to the patient, do not respond to treatment. The resistant strains, thus far, respond to alternate forms of antibiotics. Clinicians armed with the proper identification of the resistant and relatively resistant strains are better able to prescribe and treat their patients, possibly averting severe complications, mortality, or outbreaks.

According to Salyers and Whitt (10) pneumococcal strains that are penicillin-resistant are still fairly rare in the United States. Presently only 12% of clinical isolates have been determined to be resistant. Yet this 1993 figure is alarming when compared to the 1989 rate of 1.5%. It is even more disarming to consider the 30-40% mortality rate in the Western world's cases of pneumococcal meningitis (7). The greatest sources of penicillin-resistant strains are Africa and Europe. 70% of reported clinical isolates from certain European countries are resistant (2). With today's unprecedented amount of intercontinental travel, the globalization of these resistant strains is inevitable.

The first reported cases of penicillin-resistant strains of *S. pneumoniae* in the United States date back to the mid 1960's (12). Since that time there has been a steadily increasing population of these strains, culminating in several regional outbreaks (13, 18). The mechanism of penicillin-resistance involves a gene change which produces altered penicillin-binding proteins (PBP) and low affinity to beta-lactam antibiotics (10). It is doubtful that the resistance mechanism is entirely explained by this one gene change, however, due to the fact that penicillin-resistant strains are often resistant to other antibiotic categories, including the non B-lactams (erythromycin, tetracycline and chloramphenicol) (4). Countries which have

reported multiply resistant strains of *S. pneumoniae* include: Spain, Greece, Hungary, Canada, South Africa, Pakistan, Brazil, France, Australia and the United States (3, 4,11,12).

SIGNIFICANCE OF THE STUDY:

The 1995 Report of the Task Force on Antibiotic Resistance of the American Society of Microbiologists gives warning to the medical and scientific communities about the alarming escalation in antibiotic resistance among several several human pathogens (1). The conclusions of the Task Force include a strong directive that a global surveillance system be established to approach this multifaceted problem. The surveillance system should be a central repository for all information about microbial antibiotic resistance and provide continuous educational information to researchers, pharmaceutical industry, educators, and clinicians world-wide.

DIAGNOSIS

One promising method for rapid and precise characterization of infective microorganisms is PCR. Determining the causal agent of bacterial illnesses by the most rapid means possible can provide physicians and veterinarians with crucial information regarding the infectious agent. Genus, species and subtype/strain information can be available within 24-48 hours utilizing recent advances in molecular biology (19). The swiftness of the results may permit highly specific prescription antibiotic intervention, rather than the empiric approach forced on physicians/veterinarians by the slow and laborious routine laboratory culture and sensitivity testing in the traditional microbiology laboratory. According to the ASM Task Force (9), ". . .the widespread use of antibiotics use is the driving force in the development of antibiotic resistance." Early treatment with organism-appropriate, optimally

dosed antibiotics will undoubtedly help slow the progress of microorganism antibiotic resistance. PCR is an excellent tool for a systematic attack on the proliferation of antibiotic resistant organisms and a key to minimizing antibiotic misuse.

TYPE OF FINGERPRINTING:

It has been established by many investigators that prokaryotic genomes contain a variety of repeated, short, interspersed DNA sequences 6, 7, 15, 16, 17). The precise function of these sequences remains elusive, but their presence is exploitable for specific microbe identification by Rep-PCR (15). Oligonucleotide primers which complement the specific repetitive extragenic palindromic elements of a given species or strain of organism will bind to the repeated sequences. Using the bacterial genomic DNA as a template, amplification of all DNA between successive primers will produce an amplification product which can be fractionated by gel electrophoresis. The resultant DNA fingerprint of the organism is strain-specific, reproducible and available in a much shorter time than is possible by conventional bacteriologic testing. Within a twenty-four hour period, this genotypic identification is reportable.

Aggressive measures to identify the organism and its source in a rapid and precise manner also permits early epidemiological studies of nosocomial infections. Patient-to-patient or health care professional-to-patient infection accounts for considerable inpatient and outpatient morbidity and mortality (1). The ability to genotypically identify and confirm the particular strain(s) may enable early intervention into an otherwise potentially dramatic escalation of the spread of the infection.

REP-PCR:

Many prokaryotic organisms possess interspersed repetitive sequences of DNA. The repeated elements are characteristically short, non-coding sequences of uncertain function in the bacterial genome (16). Regardless of their function (s), they are quite useful in a rapid and reliable method for genotypic identification of bacteria. Primers designed to complement the element may be attached to the DNA template and DNA between successive primers may be amplified by PCR. By gel electrophoresis the resultant bands are separated, resulting in a DNA fingerprint.

MATERIALS AND METHODS:

THE ORGANISMS

Sreptococcus pneumoniae on chocolate agar stab cultures were obtained from Dr. Hassam Namduri of Clinical Laboratories, Scranton, Pennsylvania. The cultures represent samples cultured from blood and cerebrospinal fluid from inpatients of several hospitals in northeast Pennsylvania. [Other samples were obtained from various Air Force installations through the efforts of Ms. Ferne McClesky, Chief of Bacteriology, Armstrong Laboratories, Brooks AFB, San Antonio, Texas. These samples were cultured from blood, eye, sputum, throat and wound infection].

The 115 samples were streaked on 5% Blood Agar (CSM, Decatur, Ga.) and grown overnight (18 hours) in a 5% CO₂ incubator at 37 C. Single colonies were removed and replated on 5% Blood Agar, and once again incubated at 37 C in CO₂.

HARVESTING AND WASHING:

Freshly cultured, 48 hour old colonies of *S. pneumoniae* were harvested using 1.5 ml. of 1 M NaCl-10mM Tris-HCl and a wet sterile cotton swab. The suspension was centrifuged for five minutes at 5000 x g at a temperature of 4 C in a refrigerated Eppendorf Centrifuge. The fluid was decanted, and 750 ul of 1 M NaCl-10 mM Tris-HCl was added and vortexed to resuspend the pelleted organisms. Samples were again centrifuged at 5000 x g for five minutes at 4 C, and 750 ul of the same buffer added. The samples were resuspended again by vortexing.

CELL LYSIS:

To the cell suspension was added 200 ul of a 10% SDS solution and 10 ul of Proteinase K (20 mg/ml) to lyse the bacterial cells. Samples were inverted for mixing and incubated in a Syva Microtrak EIA dry bath at 37 C for one hour.

DNA PURIFICATION AND EXTRACTION:

The DNA was purified by a standard Phenol/Chloroform extraction, followed by a Chloroform extraction. 750 ul of Phenol/Chloroform was added to the sample and mixed by gentle inversion. The mixture was spun down for 10 minutes in an Eppendorf Minicentrifuge, followed by careful aspiration of the aqueous phase into a clean Eppendorf tube. This procedure was repeated using 750 ul of Chloroform. The DNA was teased out of solution by gentle inversion in 100 ul of 95% Isopropyl alcohol. The samples were refrigerated for one hour to get maximal precipitation, and centrifuged at maximal speed in the Eppendorf Minicentrifuge. The fluid was expelled and pellets allowed to air dry for 30 minutes. Following evaporation of the alcohol, 100 ul of distilled water was added, gently inverted and refrigerated for an hour.

QUANTIFICATION OF THE DNA:

The concentration and purity of the DNA was determined using the GeneQuant RNA/DNA Calculator (Pharmacia Biotech Inc.) Samples were diluted with sterilized, distilled water to bring each to a final concentration of 200 ng.

PREPARATION AND REP-PCR:

The PCR Reaction mixture was prepared according to the following formula: (Lupski, unpublished protocol)

10 ul.	5X PCR Buffer
8.875 ul.	Sterile distilled water
5.0 ul.	Dimethyl Sulfonamide (DMSO)
3.125 ul.	Nucleotide Mixture
2.2 ul.	BOX A1R Primer
0.8 ul.	TAQ Polymerase

30 ul. of the PCR Mixture and 20 ul of sample was added to the reaction vessel. The 'negative control' was 20 ul. of sterilized distilled water and 30 ul of PCR Reaction Mix. The 50 ul samples were aliquoted into the cuvettes and placed into a prewarmed (80 C) Perkin Elmer GeneAmp PCR System (9600).

The GeneAmp System was programed as follows: (Lupski protocol)

Cycle #1	95 C	7 minutes (denaturation)
Cycle #2 (X32)	90 C	30 seconds (denaturation)
	52 C	1 minute (annealing)
	65 C	8 minutes (elongation)
Cycle #3	65 C	16 minutes (elongation)
Cycle #4	4 C	Unlimited (hold)

The approximately seven hour process produced DNA amplicon, which was then loaded into a 1.5% agarose gel (Seakem) to which had been added one drop of ethidium bromide. Using both a 100 base pair and a 1 Kilobase pair ladder, the gel was subjected to 65 volts for approximately 2 hours.

RESULTS:

Due to the limited amount of time to conduct this study, no conclusive results are yet reportable. It is the intent of this investigator to seek an extension to complete and expand upon the project. Several samples were successfully fingerprinted and many aspects of the culturing, harvesting, isolation of DNA, and Rep-PCR with BOXA1R Primer were optimized during the eight weeks of the summer project. Analysis of the 115 fingerprints was not possible until all specimens are fingerprinted however.

OPTIMIZATION OF HARVEST:

During the first two weeks of the project various methods of culturing, isolating and harvesting the organism were tried. Chocolate and blood agar, and Todd-Hewitt Broth were tested for superior rates and amounts of growth with the pneumococcus. Ultimately, blood agar proved to be an excellent medium for this project. Both aerobic and anaerobic incubations were tried. The aerobic incubator supplemented with 5% CO₂ provided ideal conditions for the pneumococcus. And lastly various growth time parameters to attain a large sample were tried. While 24 hours growth is adequate for Rep-PCR, the samples grown for 48 hours were selected due to the richness of the plate growth.

HARVESTING STRATEGIES:

Glass rods bent into a 'hockey stick' shape were compared to wet, sterile cotton swabs as tools to dislodge the organisms from the blood agar plates. Without question the sterile cotton swabs immersed in TE Buffer were superior in yield of organism. Various amounts of TE Buffer were tried, ranging from 200 ul to 1.5 ml. Due to the absorbency of the 48 hour old blood agar, 1.2 ml was shown to be ideal for harvesting a total of 750 ul of organism in suspension.

ISOLATING THE DNA

The first method used to isolate the DNA was a *Staphylococcus* Protocol. In this process the sample is washed once in TE Buffer, and 10 ul of lysostaphin is added to lyse the cells (37 C for 1 hour). Subsequently it was determined that lysostaphin is not necessary. Substituting 1 M NaCl-10mM Tris-HCl and centrifuging at 4 C followed by an hour at 37 C in 10% SDS and Proteinase K, obviates the need for enzymatic lysis.

PRELIMINARY CONCLUSIONS:

Figure #1 displays a chart representing a visual comparison of the first group of *S. pneumoniae* Rep-PCR fingerprints. There appear to be seven types among the first samples:

FIGURE #1

[TYPES]

	A	B	C-1	C-2	D	E	F	G
[SAMPLE #]	2	1	7	10	6	18	26	5
	3	4	9	17		20	8	
		16	11	31		14		
			24	33		15?		
			22?			23?		

The remaining samples were too indistinct to analyze, and should be repeated [12, 21, 25, 27, 28, 29, 30, and 32]. Most of the fingerprints were too light to positively characterize. Despite the trials of 5 ul, 10 ul and 15 ul of amplicon in the 1.5% agarose gel, the bands were often faint and indistinct. Type C was divided into two subsets due to the presence or absence of a faint band in the 700 base pair region. Type C-1 showed a very faint 700 bp band while type C-2 did not.

Because there are many more samples within this outbreak which need to be identified by the Rep-PCR Fingerprint method, there may be many more types than represented here. This should not be considered indicative of the complete range of groups of this large sample.

DISCUSSION:

Rep-PCR is a relatively simple but accurate method of estimating relative degrees of similarity between bacterial samples from different sources. It is a technique which can rapidly determine possible clonal relationships between isolates, such as the *S. pneumoniae* specimens gathered for this study. Although not conclusive from the preliminary results of this study, it may also be possible to genotypically determine the three grades of penicillin susceptibility of this frequently resistant organism. The technique has the advantage of relative ease of procedure, repeatability of results and cost-effectiveness once the basic equipment has been secured.

The CDC has recently reported that in 1992 an estimated 110 million courses of antimicrobial therapy were prescribed by U.S. physicians. Many of these prescriptions were for otitis media of children. Pneumococci, the most common cause of childhood middle ear infections, have developed resistance to first-line antibiotics (penicillin, etc.), physicians are forced to use the newer, broader-spectrum antibiotics with increasing frequency. The common uses of more and newer antibiotics by clinicians introduces a rapidly evolving bacterial world

to new opportunities to become resistant. Coupled with the fact that nearly half of the total production of antibiotics in this country is directed for use in farm animals, overuse of antibiotic therapy becomes the driving force for the emergence of new cases of antibiotic resistance among microbes (9).

According to the ASM Task Force, "The laws of evolution dictate that microbes will eventually develop resistance to nearly every antibiotic"(9). It is time for scientists, clinicians, and the pharmaceutical industry to reassess their faith in the absolute power of antibiotics. It is also time for all clinical laboratories to be equipped and for personnel to be trained in molecular strategies for organismal identification, including Rep-PCR.

REFERENCES CITED

- (1) Emori, T.G. and Gaynes, G.P. (1993). An Overview of Nosocomial Infections, Including the Role of the Microbiology Laboratory. *Clin. Micro. Rev.* 6:428-42.
- (2) Hermans, P.W.M. et al. (1995). Comparative Study of Five Different DNA Fingerprint Techniques for Molecular Typing of *Streptococcus pneumoniae* Strains. *33:1606-12.*
- (3) Jette, L.P. et al. (1994). Pneumococcal Resistance to Antimicrobial Agents in the Province of Quebec, Canada. *J. Clin. Micro.* 32:2572-75.
- (4) Kanavaki, S. et al. (1994). Antibiotic Resistance of Clinical Isolates of *Streptococcus pneumoniae* in Greece. *J. Clin. Micro.* 32:3056-58.
- (5) Krisher, K.K. and Linscott, A. (1994). Comparison of three Commercial MIC System, E Test, Fastidious Antimicrobial Susceptibility Panel, and FOX Fastidious Panel, for Confirmation of Penicillin and Cephalosporin Resistance in *Streptococcus pneumoniae*. *J. Clin. Micro.* 32:2242-45.
- (6) Lefevre, J.C. et al. (1993). DNA Fingerprinting of *Streptococcus pneumoniae* Strains by Pulsed-Field Gel Electrophoresis. *J. Clin. Micro.* 31:2724-28.
- (7) Lupski, J.R. and Weinstock, G.M. (1992). Short, Interspersed Repetitive DNA Sequences in Prokaryotic Genomes. *J. Bacteriol.* 174:4525-9.
- (8) Nolte, F.S. et al. (1995). Detection of Penicillin-Resistant *Streptococcus pneumoniae* with Commercially Available Broth Dilution Panels. *J. Clin. Micro.* 33: 1804-6.
- (9) Report of the ASM Task Force on Antibiotic Resistance. (1995). American Society for Micro. Wash. DC.
- (10) Salyers, A.A. and Whitt, D.D. (1994). *Bacterial Pathogenesis: A Molecular Approach.* ASM Press, Wash. D.C. 322-29.
- (11) Sessegolo, J.F. et al. (1994). Distribution of Serotypes and Antimicrobial Resistance of *Streptococcus pneumoniae* Strains Isolated in Brazil from 1988 to 1992. *J. Clin. Micro.* 32:906-11.
- (12) Sibold, C. et al. (1992). *Genetic Relationships of Penicillin-Susceptible and -Resistant Streptococcus pneumoniae Strains Isolated on Different Continents.* *Inf. and Immun.* 60:4119-26.
- (13) Versalovic, J. et al. (1993). Penicillin-Resistant *Streptococcus pneumoniae* Strains Recovered in Houston: Identification and Molecular Characterization of Multiple Clones. *J. Inf. Dis.* 167:850-6.

- (14) Versalovic, J. et al. (1993). DNA-Based Identification and Epidemiologic Typing of Bacterial Pathogens. *Arch. Path. and Lab. Med.* 117: 1088-98.
- (15) Versalovic, J. et al. (1994). Genomic Fingerprinting of Bacteria Using Repetitive Sequence-Based Polymerase Chain Reaction. *Meth. in Mol. and Cell. Biol.* 5:25-40.
- (16) Vesalovic, J., Koeuth, T. and Lupski, J. (1991). Distribution of Repetitive DNA Sequences in Eubacteria and Application to Fingerprinting of Bacterial Genomes. *Nuc. Ac. Res.* 19:6823-31.
- (17) Versalovic, J. and Lupski, J.R. (1995). Differential Subsequence Conservation of Interspersed Repetitive *Streptococcus pneumoniae* BOX Elements in Diverse Bacteria. (Unpublished).
- (18) Virolainen, A. et al. (1994). Comparison of PCR Assay with Bacterial Culture for Detecting *Streptococcus pneumoniae* in Middle Ear Fluid of Children with Acute Otitis Media. *J. Clin. Micro.* 32:2667-70.
- (19) Woods, C.R. et al. (1993). Whole-Cell Repetitive Element Sequence-Based Polymerase Chain Reaction Allows Rapid Assessment of Clonal Relationships of Bacterial Isolates. *J. Clin. Micro.* 31:1927-31.

A preliminary report of the Bioacoustic and Biocommunication
Division (Armstrong Laboratory)
HRTF system and its comparison with free field
listening conditions

Alan D. Musicant
Assistant Professor
Dept. of Psychology

Middle Tennessee State University
Box X063
Murfreesboro, TN 37132

Final Report for:
Summer Faculty Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling AFB Washington, DC

and

Armstrong Laboratory
Bioacoustics and
Biocommunications Division

September 1995

A preliminary report of the Bioacoustic and Biocommunication
Division (Armstrong Laboratory)
HRTF system and its comparison with free field
listening conditions

Alan D. Musicant
Assistant Professor
Dept. of Psychology
Middle Tennessee State University

Abstract

A methodology for recording and playback of Head Related Transfer Functions (HRTFs) was established for use with the apparatus developed by the Bioacoustics and Biocommunications Division of the Armstrong Laboratory at Wright Patterson Air Force Base (AL/CFBA). The process involved the introduction and testing of a new microphone recording system, a paradigm for negating the effects of listening over headphones and an experimental design for testing the efficacy of the recorded HRTFs. Additionally, an extended frequency range for the sound stimuli was introduced.

A preliminary report of the Bioacoustic and Biocommunication
Division (Armstrong Laboratory)
HRTF system and its comparison with free field
listening conditions

Alan D. Musicant

Introduction

There has been a tremendous increased interest in the area of sound source localization over the last decade. A major focus of this interest has been aimed at establishing recording techniques so that headphone simulations of free-field sound sources could be established. The interest has gained impetus because of the potential applications tied to communications and entertainment. All of the procedures have been designed to accurately determine the contribution of the pinnae to performance in locating the position of a sound source. This contribution takes the form of an enhanced ability to locate a sound in the vertical dimension as well as with front-rear discrimination. A corresponding interest in determining Head Related Transfer Functions (HRTFs) for use in playback, over headphones, of digitally synthesized signals has also occurred. HRTFs incorporate pinna, head and torso spectral contributions as well as interaural cues to sound localization. Use of HRTFs and synthesized signals containing spatial information has practical applications in the areas of cockpit communications, localization in noisy environments and also for areas in entertainment.

HRTFs have been determined in a variety of ways over the last two decades. The most commonly used methods have involved insertion of a miniature microphone into the ear canal of a human subject, the use of various acoustic mannequins and the use of larger recording microphones in dummy heads. These techniques have utilized both blocked and open ear canals. Stimuli have been either click trains or gaussian noise bursts delivered from a variety of sound source locations. Generally, all source locations have been at the same distance.

The Bioacoustics and Biocommunication Division, Armstrong Laboratory (AL/CFBA), Wright-Patterson Air Force Base developed a sound localization laboratory several years ago. This facility utilized a large anechoic chamber and a geodesic dome with 272 loudspeakers. The loudspeakers were arrayed at approximately 15° intervals in both azimuth and elevation. The personnel at AL/CFBA developed a technique that utilized an acoustic mannequin of local design, a mold of subjects' ears, and the use of a 1/2"

microphone (B&K). Localization of sound sources under headphone listening conditions, as compared to free field listening, were compared for the horizontal plane (0° elevation) and the median saggital plane (0° azimuth). Results were comparable.

A different (improved) recording system and a systematic evaluation of the results was the goal for the summer research grant. Towards this end the development and testing of a miniature microphone recording system and the playback system were the major goals. Additionally, results were to be compared for the full range of sound source locations, i.e., 360° azimuth at elevations ranging from -45° to +90°.

Methods

A miniature electret microphone (Knowles Electronics, model EA1934) was chosen as the microphone with the best frequency characteristics and the size necessary to place in molds of an ear canal. This microphone has been used by a number of other laboratories.

Ear molds were made from subjects selected from a subject pool maintained at the lab. These molds consisted, in their final form of a silastic material with the density of human skin. The ear mold was complete down to approximately 8 mm into the ear canal. These molds were utilized for the recording of the sound spectrum from the various loudspeakers and the interaction of that sound with the pinna, shoulders and torso (HRTF) of a human sized mannequin. Pink noise was delivered to each loudspeaker individually. The sound emanating from each loudspeaker was recorded by the miniature microphone (Knowles Electronics, model EA1934) embedded into a clay material on the surface of the acoustic mannequin head. The microphones were placed so that, when the ear molds were attached to the mannequin head, they projected into the ear canal part of the mold. A small hole was cut into the end of the mold ear canal to facilitate this placement. The microphone was located at what would be, in a person, approximately a 5mm depth at the end of a blocked ear canal. This is the position that Blauert has recommended as the most appropriate location to record HRTFs that bear a linear relationship (for all spatial locations) to that which is recorded at the eardrum.

The basic technique of digitally processing the spectra of a sound source has been well established. The use of the Fast Fourier Transform (FFT) in this effort is common. These methods delineate the spectral transformations that occur as a function of the sound source location and the interaction with the pinna, head and torso. These method can also be used to establish another important parameter of the localization process, i.e. the interaural arrival time difference (ITD). Many laboratories have, however, chosen to determine the ITDs for different locations utilizing other methodologies. These include minimum phase or

linear phase assumptions for the transformation. The personnel at AL/CFBA had previously utilized a method of ITD determination that relied upon the difference in peak arrival time for a sine wave. While reliable, this was a time consuming process that depended upon visual determination of the peak arrival at the two recording microphones. A project was underway that was designed to utilize a linear phase assumption for frequencies between 100 and 5000 Hz. These phase calculations could then be used to determine an ITD for sound originating from any location on the geodesic dome.

Recordings were made for three sets of ear molds. Playback and localization judgments were to be made with a comparison to free field localization performance. One additional problem to solve prior to the evaluation of the headphone listening condition was the issue of how to equalize for the headphone transformation of the HRTFs. This was necessary since the sound emanating from the headphones would interact with the pinna, creating an additional transformation of the HRTFs. The solution to the problem is complex but at least one possibility offering a solution was to record the transformation of a pink noise delivered over the same set of earphones that were going to be used in the actual experiment. The headphones (Sennheiser 520) were placed over the mold pinnae that were mounted onto the head of the acoustic mannequin. The headphones were adjusted so that they appeared to sit on the head in a location that corresponded with a "comfortable fit" as judged by one of the subjects. A pink noise signal was delivered over the headphones and the resulting waveform, transformed by the electronics, the headphones and the pinna, was recorded. The FFT of this waveform was then inverted and applied to the HRTFs that had been recorded earlier. Previous analysis of different headphone positions had indicated that small changes in earphone position, when removing and refitting the headset, had only small effects on the transformation of spectra when using the gaussian noise stimulus.

The procedures outlined above were utilized to determine HRTFs of subjects and to insure that the sounds delivered over headphones were free of any artifact that might interfere with the spatial localization performance of subjects while listening over headphones. An outline of the basic procedure and the components at each step are further delineated below.

Measurements or procedure:	Components present
1. Determine spectra of sound from each loudspeaker	Microphone characteristics Loudspeaker characteristics Electronics signatures
2. Place miniature microphones in ear molds mounted on acoustic mannequin	HRTF Microphone characteristics Loudspeaker characteristics Electronics signatures
3. Divide measurement 2 by measure 1 (this is equivalent to subtracting the two measures)	HRTF
4. Sound played over headphones to subjects	HRTF headphone characteristics pinna effects for the interaction of sound from headphones electronics signatures
5. Determine spectra of sound delivered over headphones mounted over ear molds on acoustic mannequin	Microphone characteristics headphone characteristics electronics signatures Pinna transform from placement of headphones
6. Determine microphone and electronics characteristics through microphone calibration	Microphone characteristics electronics signature

The manipulations of the digitally determined spectra is as follows:

step (2) - step (1) = HRTF

Simply playing back these HRTFs over headphones results in some undesirable characteristics as outlined in step (4) above. therefore

(4) - (5) + (6) = HRTF

This sequence results in a signal delivered to the subject that contain only the information from the HRTF associated with a particular spatial location and is devoid of all unwanted characteristics.

Outlined in the following table is the experimental design. Six experimental listening conditions, three under binaural listening conditions and three under monaural listening conditions will be utilized to test the efficacy of the HRTFs and the various procedures designed to deliver the signal to a headset. All subjects will participate in all listening conditions.

	1	2	3
Binaural	Free field	live listening over headphones	HRTF listening
Monaural	Free field	live listening over headphones	HRTF listening

Condition 2, live listening over headphones, is an intermediate condition between the free field listening and the HRTF listening condition. This condition involves listening to a pink noise signal presented from the loudspeakers in the geodesic dome that is picked up by the same miniature microphones utilized in the recording of the HRTFs. The sound will be delivered via headphones to the listener, who will be situated in another room.

This sound will be compensated for via the transformation technique used to compensate for the problem with HRTF delivery over headphones. There is no synthesis of HRTFs required for this step. A similar arrangement was used by Batteau (1968), who established the possibility of using model pinna in different environments.

We will be comparing the performance data for localization for the six listening conditions. Monaural listening conditions are a requirement since it is possible that, under some conditions, monaural delivery of sound may occur. These experiments will be carried out by the personnel of AL/CFBA during the fall of 1995 and the winter of 1996. Analysis of data will require determination of errors (relative to the actual location of the present sound source) in all conditions. Comparisons of performance will utilize an analysis of variance model for a same subjects design. Additionally, multiple post-hoc tests will be conducted where appropriate.

References

Blauert, J. (1982) Spatial Hearing. The psychophysics of human sound localization. MIT Press:Cambridge,MA.

THE EFFECTS OF EXTREME ENVIRONMENTS ON HL-60 PROLIFERATION AND CYTOKINE
PRODUCTION

Edward H. Piepmeier, Jr.
Assistant Professor
College of Pharmacy

University of South Carolina
Columbia, SC 29208

Final Report for:
Summer Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington D. C.

August 1995

THE EFFECTS OF EXTREME ENVIRONMENTS ON HL-60 PROLIFERATION AND CYTOKINE PRODUCTION

Edward H. Piepmeier, Jr.
Assistant Professor
College of Pharmacy
University of South Carolina

Abstract

The primary goal of this summer's research was to identify how hypobaric and hyperbaric exposure alters the proliferation and cytokine production of leukocytes. In addition, these cells will be tested in the microgravity environment aboard the STS-69 Space Shuttle Flight. HL-60 cells, an immortalized human promyelocytic cell line, were used as a model for this study. While these cells could be expected to exhibit a response different than normal human promyelocytic cells, the variation associated with HL-60 cells is much less. We found that exposure to rapid decompression resulted in a significant increase in cell proliferation, while exposure to a high pressure resulted a proliferation rate lower than but not significantly different than the control cells. Corresponding changes were observed in IL-1, IL-2 and nitrite production. These preliminary results indicate that leukocytes may be a good marker for decompression sickness and hyperbaric oxygenation therapy which is used to treat the disease. Further studies should be conducted to examine the sequential effects of rapid decompression and hyperbaric oxygenation on the cells.

THE EFFECTS OF EXTREME ENVIRONMENTS ON HL-60 PROLIFERATION AND CYTOKINE
PRODUCTION

Edward H. Piepmeier, Jr.

INTRODUCTION

This study was conducted to further elucidate the effects of hyperbaric oxygenation exposure on leukocytes. In addition, the effects of rapid decompression were studied for the potential development of a positive control for the study of leukocytes in a diseased environment. Cell proliferation rate (table 1) and the cytokines nitric oxide (table 2), interleukin-1 (table 3) and interleukin-2 (table 4) were monitored.

The overall goal was to elucidate the effects of hyperbaric oxygenation therapy on potential targets for improved response in treating disease states. Hyperbaric oxygenation therapy has been approved for use in: air or gas embolism, carbon monoxide poisoning and smoke inhalation, carbon monoxide poisoning complicated by cyanide poisoning, clostridial myonecrosis, crush injury, compartment syndrome, and other acute traumatic ischemias, decompression sickness, enhancement of healing in selected problem wounds, exceptional blood loss anemia, necrotizing soft tissue infections, refractory osteomyelitis, radiation soft tissue damage and osteoradionecrosis, compromised skin grafts and flaps, and thermal burns. The areas currently under investigation for hyperbaric oxygen treatment include; blunt force injuries, penetrating injuries, open fractures and bone grafting, chemical thermal burns, ionizing/nonionizing radiation injuries, skin grafts and flaps in compromised tissue, nerve regeneration, aseptic bone necrosis, organophosphate poisoning, septic shock, cyanide poisoning, cerebral edema, cold-frostbite, spinal cord injury and explosive decompression. However, like aspirin which was used for thousands of years without knowing its mechanism of

action, the mechanism of action for hyperbaric oxygen therapy in many of these disease states is unclear. Only twenty years ago aspirin was identified as a cyclooxygenase inhibitor and it has only been within the past few years that some of the immune effects of hyperbaric oxygenation therapy have been identified. This study further defines the immune effects of hyperbaric oxygenation treatment which are responsible for the therapeutic use in the treatment of many of the above indicated disease states as well as the effects of rapid decompression. In addition, results obtained in support of these preliminary studies could be used to identify the potential of hyperbaric oxygenation therapy in other disease states.

METHODS

Cell Preparations - Sterile cell culture techniques were followed. HL-60 cells were cultured in RPMI medium supplemented with 1% penicillin-streptomycin, 1% sodium pyruvate, 1% non essential amino acids and 10% heat inactivated fetal calf serum.

Nitric Oxide Assay - Nitric oxide content in solutions is determined with the Griess Reagents (0.8% sulfanilic acid and 0.5% N,N,-dimethyl-alpha-naphthylamine, both in 5N acetic acid) (Baxter Medical Corp., Sacramento, CA.) The assay colorimetrically measures nitrite concentration which correlates with Nitric Oxide. In a typical experiment, 10,000 cells/100 ul are placed in a well of a 96 well microtiter plate with 100 microliters of each Griess Reagent. The plate is incubated for 10 min at 37 degrees C. Absorbance is read at 546 nm on a SLT Spectra Shell microplate reader.

Interleukins Assay - Interleukins were assayed using commercially available 96 well plate immunoassays.

RESULTS

This study examined the effects of hypobaric and hyperbaric exposure on HL-60 cells. Changes in response of the mitogenic response of HL-60 cells (table 1) and blood nitrite concentrations (table 2) interleukin-1 concentrations (table 3) and interleukin-2 concentrations (table 4) were monitored. Nitrite concentrations were determined colorimetrically at 570 nm after incubation of the mixture for ten minutes at 37C. Cell proliferation was determined manually with a hemacytometer and all counts are based upon one hundred microliter samples which were stained with trypan blue to assess cell viability.

Table 1. Proliferation rates. Increase in number of cells per milliliter from one day following treatment until three days following the treatment (standard error)

	x10 ⁴
Control	46.30 (35.02)
Hyperbaric	39.15 (11.93)
Decompression	139.90 (41.01)

Table 2. Nitrite Measurements (mM)

Experiment	# 1	# 2	Avg
Control	0.304	0.314	0.309
Hyperbaric	0.281	0.291	0.286
Decompression	na	0.336	0.336

The control cells and the cells exposed to hyperbarics or decompression all gave produced nitrite proportional to their proliferation rate (see table 2)

Table 3. Interleukin-1 values observed in supernatants from each of the three treatments. (standard deviation)

Control	0.284 (30pg/ml)
HBO	0.291 (30pg/ml)
Altitude	0.333 (30 pg/ml)

Table 4. Interleukin-2 values observed in supernatants from each of the three treatments. (standard deviation)

Control	0.935 (665 pg/ml)
HBO	0
Altitude	0

Interleukin-1 is linked to host responses to injury and infection, but prolonged IL-1 production is linked to inflammatory or autoimmune diseases. IL-1 binding fragment (which inactivates IL-1) was initially successfully used to treat animals for shock. However clinical trials demonstrated that the drug was ineffective and possibly detrimental to humans.

Based upon previous studies , because interleukin-1 is linked to the inflammatory response, it is hypothesized that our readings will be lower for the treatment group which is exposed to hyperbaric treatment and that it will be significantly higher in the treatment group which is exposed to altitude.

IL-1 can possibly lead to increased IL-2 expression and increased differentiation. IL-2 has been isolated from peripheral blood mononuclear cells and Jurkat cells (human leukemia). IL-2 is produced when these cells are stimulated by some mitogens and has demonstrated activity to further stimulate the antitumor action of T-cells.

CONCLUSIONS

These studies demonstrate the effect of hypobaric and hyperbaric exposure on leukocytes. Of particular interest is the effect of hypobaric exposure on leukocytes. The leukocytes proliferate similar to what would be seen in an inflammatory response. In addition, the leukocytes also expressed elevated amounts of IL-1 when exposed to rapid decompression relative to control and leukocytes exposed to hyperbaric oxygenation. IL-2 was non-existent in the supernatants of the treated groups which demonstrates that IL-2 expression is not increased by the treatments.

Continued studies must be made into the effects of hyperbaric oxygenation on the leukocytes which have undergone decompression.

PROCESS CONTROL OF HYBRID OXYGEN SYSTEMS

Randel Miles Price
Assistant Professor
Department of Chemical Engineering

University of Mississippi
University, MS 38677

Final Report for
Summer Faculty Research Program
Armstrong Laboratory

Sponsored by
Air Force Office of Scientific Research
Bolling Air Force Base, DC
and
Armstrong Laboratory

September 1995

PROCESS CONTROL OF HYBRID OXYGEN SYSTEMS

Randel M. Price
Assistant Professor
Department of Chemical Engineering
University of Mississippi

Abstract

Hybrid oxygen systems were reviewed to identify feasible improvements in the process control system. Suggested improvements should reduce the average rate of consumption of refrigerant air.

This reduction should, in turn, reduce the operating cost of the systems. Other recommendations should improve the flexibility and reliability of the systems and increase the value of the information gathered from the proof-of-concept unit.

PROCESS CONTROL OF HYBRID OXYGEN SYSTEMS

Randel M. Price

The Oxygen Systems Function of the Crew Technology Division, Armstrong Laboratory, Brooks Air Force Base, is engaged in a program to develop "hybrid oxygen systems". These combine a pressure swing adsorption molecular sieve oxygen concentrator (often abbreviated as OBOGS, on-board oxygen generating system, or MSOGS, molecular sieve oxygen generating system) with an oxygen liquefier to permit storage of purified oxygen. The goal is to reduce the size and weight of aircraft oxygen systems and reduce the logistics complications imposed by ground crew handling and storage of LOX [1].

A hybrid oxygen system (HOS) test facility has been installed at Brooks AFB. This system uses two helium cold heads to provide cooling, and is designed to produce about 40 slpm of 95% oxygen [2]. Test runs on this unit are scheduled to begin in late summer 1995 [1].

Existing and proposed designs and operating plans for various facets of the hybrid oxygen system program have been reviewed with the object of identifying opportunities for improved flexibility, performance, reliability, and savings through applications of process control. This report summarizes potential concerns and modifications in two categories:

- 1) HOS -- using existing equipment only
- 2) HOS -- with instrumentation changes

The suggestions have been developed conceptually, but no significant design work has been done. In many cases, the suggestions made for one phase of the project are also applicable to subsequent categories of the design as well.

In all of these systems, the raw materials are free and operating labor is minimal, so operating costs are essentially limited to the energy required to produce compressed air and for providing cooling. Consequently techniques for reducing air consumption or cooling load are a primary motive for the recommended changes.

1.0 HOS -- existing equipment

This section suggests areas for investigation and enhancement of the computer control displays which can be implemented without any changes to the HOS system hardware.

1.1 Computer Control Displays

Effective operator displays tend to group related information so that it is readily interpreted, instead of simply displaying arrays of numbers. A strict numerical list of temperatures or pressures is probably less efficient than a grouping arranged to show cooling effectiveness or pressure drop. For example, many process plant control displays are schematic process flow diagrams of a plant or a section of the plant. Individual streams (lines) and vessels are flagged with the flow, pressure, and temperature information. This way the operator can see the relevant information without having to search through a list of data points.

The HOS system is readily divided into two portions: the refrigerant, or air, side of the system is independent of the oxygen side. The operating modes for the refrigerant side (cooldown, deriming, warm-up) are independent of what is happening on the oxygen side (startup, fill, draw, purge), while the fill and run modes also cycle between dewars. Consequently, it seems reasonable to separate the two sides within the display.

1.2 Cooldown

The extent of frost formation problems during the cooldown stage should be qualified and quantified during the HOS test runs. Future designs will benefit, since frost protection will need to be provided (as expander bypasses, etc.) Besides temperature and pressure measurements, it might be interesting to try to calculate and record an approximate cooling load as $(FM1 \text{ mass})(DTS1 - DTS4)$ for the cold heads and $(FM1 \text{ mass})(DTS4 - DTS5/6)$ for the JT valve. This would indicate the relative contributions of the different sections of the cooling loop, and the total could eventually be used as a target for control, since cooling load is of greater importance than the coolant flow or individual temperatures.

During initial cooldown, the DP across the JT (P7 - P9) valve is an observed variable, used to determine when the valve must be adjusted. Consequently, it would be useful to have it calculated and displayed on screen.

1.3 Deriming

Similarly, once the deriming cycle begins, DP through the system is an indicator of the need for deriming, and it would be useful to calculate the value and display it on the monitor. The switching of flow direction should not cause problems if the absolute value of P3-P4 is used. Does it make sense to let the FM1 setpoint drift as the run progresses, or should an effort be made to hold it at the initial value? The ADL data [2] seem to show the flow rate drifting slightly downward, although the upper bound seems to stay fairly steady. A drift in the flow rate is inevitable if the deriming cycle is not entirely effective. This might occur if the lower pressure outlet flow is much colder than the high pressure coolant inlet, so that the return stream is not capable of absorbing all the material deposited [3]. In this case, the flow rate will necessarily drift; consequently, trying to maintain a fixed setpoint may cause the deriming cycle to oscillate ineffectively. On the other hand, it may be that the ADL test run switched before deriming was complete, and that a fixed setpoint is viable. HOS tests are needed to clarify this matter.

It might be interesting to track the deriming cycle time interval. ADL reports that a 5 minute cycle is "necessary and adequate" [2]. Cycle length will likely vary with the quality of the inlet coolant. The HOS test rig should have drier air than ADL used, so a longer cycle seems reasonable. As the rig is put through its paces, the cycle time might be a good indicator of variation in feed gas quality.

At some point, it may be desirable to run with a fixed cycle time, using a timer to reverse flow. This would be the simplest way to automate the process, but could not compensate for variations in feed gas conditions, cooling load, etc.

1.4 Oxygen Production

On startup, the concentrator should be run without taking product until the oxygen concentration reaches the desired level [4]. This uses the entire air flow as purge and brings the system to running conditions faster.

During normal operation, flow through the concentrator can be adjusted to control product composition. If oxygen concentration drops too low, the operator can back off on FC2 (oxygen to liquefier) to shift to a more favorable location on the concentrator flow-composition curve. An automatic control solution is described in Section 2.5.

The dewar mass balances can be calculated on-line and displayed at the console. This would provide a check on the dP cells. The mass balance for filling dewar 1 can be calculated and displayed as $M_1 = M_{1(0)} + \int (FC\ 2 - FM\ 4)dt$, while for the draw mode it is $M_1 = M_{1(0)} + \int (-FC\ 1)dt$. The equivalent balances for dewar 2 are

$M_2 = M_{2(0)} + \int (FC\ 2 - FM\ 3)dt$ and $M_2 = M_{2(0)} + \int (-FC\ 1)dt$. In these equations all flow meter readings are assumed to have units of mass per time. The values subscripted with a zero are initial values at the time balance information begins. A simple explicit Euler integration algorithm seems more than adequate for on-line evaluation of these balances.

There will probably be a lag between heater shutoff and a decrease in the rate of pressurization. Shutting off the heater before the target pressure is reached should help reduce the chance of over pressurizing and thus unnecessary venting of oxygen, so whenever dewar pressure (P14 or P15) exceeds 4.9 atm (72 psia), the corresponding heater (1 or 2, respectively) should be off. The shutoff point will need to be adjusted if the actual operating pressure is not 5 atm and to reflect experience with the system (4.9 atm is chosen without quantitative support).

1.5 Dewar Switching Lockouts

Normally, the dewars will be filled at maximum rate, subject to concentrator performance. Once the dewars are full, the oxygen flow can be reduced to minimum. If the time elapsed between fill intervals is too long, so that substantial time is required to cool the system back to operating conditions, it might be worth filling at a slower rate.

A full dewar that has started to self-pressurize should not be switched back to fill mode. If this is done, the dewar will blow through vent line regulator RA3/RA4 until the dewar pressure drops from about 5 atm to about 3 atm. Unless the vented oxygen is recovered (Section 2.9), it will be wasted. It is less costly, although still inadvisable, to switch to draw mode a dewar that is already empty. This will pressurize the dewar and reduce the LOX make.

When both dewars are full, it will be necessary to block out the concentrator by closing OSV1 and/or setting FC2 to zero. If LOX is not being made, the only load on the coolant

loop is the mass of the exchangers, so it may be desirable to minimize the work done in the coolant loop (Section 2.2).

When both dewars are full, dewar level could be forced down by increasing heater load (to raise pressure and force pressure relief or provide for faster draw) or by bleeding oxygen product through OSV3. The present system is not really designed for storage of LOX for purposes other than peak shaving of demand, a topic that is discussed further in Section 2.6.1.

If both dewars are empty, the draw must be shut down by closing OSV2 and OSV3 and/or setting FC1 to zero. A warning (no oxygen available) should be issued and FC2 should be set to the maximum value.

Dewar level control may be noisy. If so, one can increase the dead band or use a delay to keep the fill/draw switching from bouncing around too much. The dead band refers to the size of the input reversal needed before the control output is changed. Increasing dead band prevents rapid back and forth oscillation of a manipulation. For instance, in a noisy system set at 50%, it may be wise to avoid taking control action until the measurement falls below 45% or above 55%, so that noise alone is insufficient to cause a control change. The measurement might also be averaged or filtered to reduce the impact of noise.

If boiling is sudden and intense, inverse response might be seen. Inverse response means that the initial response is in an opposite direction to the eventual steady-state response. In a liquid system, an increase in boiling may initially create enough turbulence to cause the apparent level to rise before increased vaporization causes it to fall.

1.6 Warm-up

Finishing every experimental run with a warm-up phase seems unnecessary. Unless cooldown performance is being measured, it may make sense simply to purge the oxygen side of the system and shutdown the refrigeration side without warming it back up. This might reduce the cooling load the next time the system is started, particularly if the vacuum, with its insulating effect, is maintained.

Eventually, it may be worthwhile to reconfigure valves to permit a small amount of LOX to be held in the dewars between runs to keep the dewars cold. A study of a large liquid hydrogen storage tank cited by Timmerhaus and Flynn [5] showed that partially filling a tank and allowing it to chill before completing filling used less than half as much gas as attempting to fill the warm tank rapidly. This suggests that steps to keep the dewars cold will reduce total energy and air consumption.

1.7 Alarms and Shutdown Interlocks

The current HOS operating plan includes eight alarm conditions [2]; if one of these limits is exceeded, the system will be shut down [6]. In three cases, the limit does not signal a hazard, but rather abnormal operation. It may be wise to review whether immediate shut down is required. Instead, it may be prudent to first take steps to confirm and remedy the problem. If the system is shut down, the exchangers will begin to warm, and the cooldown phase must be at least partially repeated. Since the cooldown phase is lengthy, and the least relevant to future work, it may be wise to avoid inessential shutdowns. This section summarizes the alarm conditions and possible responses. The first five conditions probably justify immediate shutdown on confirmation.

Low temperature at DTS10 signals LOX in the product line to cabin plenum. Some confirmation might be obtained by checking the dewar level and temperature to see if dewar is full. If dewar is empty, it is less likely that substantial liquid is in line. If confirmed, the system probably should be shut down.

A high reading on P14 or P15 signals high dewar pressure. In this case, the heaters should be shut off to reduce the rate of pressurization, while the vent valve (OSV4/5) can be opened to bleed off oxygen to the atmosphere. If the condition persists, shut down the system.

A high reading on P16 signals vacuum vessel over pressure -- pressure relief should open automatically if the pressure continues to increase. This condition suggests immediate shutdown.

TC6 monitors the temperature in the drip pan at the bottom of the apparatus. If it shows a drop in temperature it may mean that liquid oxygen is spilling from the dewars. The system should be shut down if this occurs.

The mass spectrometer sample from NV5 monitors product oxygen content. If it is low, it may mean that oxygen is being diluted through leakage from the system, so it should be compared with a sample from NV2 to see that oxygen is entering the liquefaction system at design concentration. If the vacuum pressure (P16) is increasing, it may be confirming a leak. In this case, shut down the system.

Low temperature at TC4 or TC5 signals LOX in dewar vent. If detected, it may mean that dewar is overflowing. For confirmation, can check dP cell output (dP1, dP2) or mass balance from flow meter readings (especially if the mass balance is being recorded on-line); a low level indication might suggest the LOX in the vent line is caused by something other than overflow. Checking dewar temperature may also be worthwhile, since if dewar is warm it is less likely that LOX in vent is caused by overflow. If overflow is not confirmed, the most reasonable actions appear to be to switch filling to the other dewar by closing dewar vent, or to shut down oxygen flow from concentrator (OSV1) to stop filling while allowing coolant cycle to continue.

A low reading on P5 signals low pressure in compressed air plenum. This pressure will probably be set somewhere around 60 psig, with the alarm set about 55 psig. Best operating pressure for the concentrator is probably around 30-40, so refrigerant loop will be most effected. Low pressure shutdown probably is not essential, since operation is feasible, although ineffective, almost to atmospheric pressure; consequently, there should be ample time to attempt remedies.

The mass spectrometer sample from NV2 monitors concentrator outlet composition, and if it is low, the operator should check supply pressure (P2) as well. If the concentration is low, OSV2 can be closed to block in the concentrator and see if the increased purge rate will bring the concentration back up within range. If the pressure is out of range, can change the setting on compressed air regulator RA1 (MANUAL) to correct. This may move the concentrator back to desired operating range. Automatic control (as described in Section 2.5) may also be effective.

2.0 HOS -- with instrumentation modifications

A number of opportunities for expanding HOS performance have been identified, but require modifications of the system. Typically, they call for installation of control valves to permit throttling and control of flow rates or pressures within the system.

2.1 Refrigerant Flow Control (for constant flow operation)

Adding a flow control valve in series with refrigerant supply valves ASV3 and 5 would allow adjusting the refrigerant flow rate to compensate for the increased pressure drop caused by frost buildup and smooth the refrigerant flow rate (Figure 2). With this configuration, it should be possible to maintain the average coolant flow at a higher rate and consequently keep the cooler load and LOX production at a slightly higher average as well.

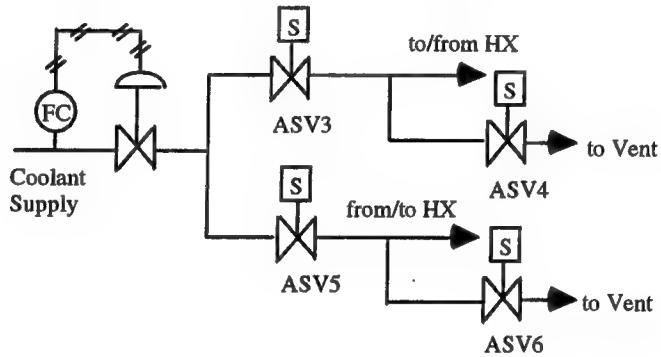


Figure 2

The deriming cycle would still be needed to eliminate ice buildup. A single flow controller would be required; either two flow control valves in line with ASV3 and ASV5 or a single valve upstream of the line split could be used. This idea depends on the coolant air supply pressure being high enough to permit throttling, as presently seems to be the case (ADL seems to use 60 psig compressed air, but the initial inlet pressure is 40 psig [2]). Note that although the bleed air specification provides a nominal pressure of 55 psig, it permits a minimum level of 20 psig [7]. At the lowest pressure, the flow controller will probably be wide open and the system will behave as if this modification had not been made, although at a slightly reduced inlet pressure (because of the additional valve pressure drop).

A preliminary sizing calculation (using ISA sizing equations [8]) indicates that a line-size globe valve with a normal operating C_v of 0.7 and a wide open C_v of 2.1 would serve this purpose. The normal operating C_v is based on the design flow of 23.7 scfm, a coolant supply pressure of 60 psig, and a system inlet pressure of 40 psig. The wide open C_v is based on 23.7 scfm, a supply pressure of 43 psig, and an inlet pressure of 40 psig. An

intermediate value, calculated at 23.7 scfm, 60 psig supply and 57 psig inlet to represent maximum frost formation, yields a C_V of 1.8.

A crude estimate of the potential for increased LOX production can be made using data from the ADL report [2]. During run 7, the average coolant air flow after the start of the deriming cycle is about 20.75 scfm or 11.76 g/s (mean value, time average is 20.73 scfm). The design rate of the system is 23.7 scfm or 13.41 g/s. If 50% of the difference can be recovered, an additional 0.83 g/s of air flow is obtained, providing an additional 8 W of cooling, which could liquefy up to 0.04 g/s of additional oxygen. This is an improvement of about 6%.

$$\frac{(20.75 \text{ cfm} \times 1 \text{ atm})}{(530 \text{ R}) \left(0.73 \frac{\text{cf} * \text{atm}}{\text{lbmol} * \text{R}} \right)} \left(28.97 \frac{\text{lb}}{\text{lbmol}} \right) \left(\frac{\text{min}}{60 \text{ sec}} \right) \left(\frac{454 \text{ g}}{\text{lb}} \right) = 11.76 \frac{\text{g}}{\text{sec}}$$

$$\left(0.83 \frac{\text{g}}{\text{s}} \right) \left(1.044 \frac{\text{J}}{\text{g K}} \right) (94.85 - 85.51 \text{ K}) = 8.09 \text{ W}$$

$$\frac{(8.09 \text{ W})}{\left(199.70 \frac{\text{J}}{\text{g}} \right)} = 0.04 \frac{\text{g}}{\text{sec}}$$

The data suggest that the mean filling rate is closer to 0.41 g/s than the design value of 0.66g/s, but it isn't clear what the target flow rate was during the run. Consequently, the above calculation is a very rough estimate.

Implementation of this control system will change the triggering of the deriming cycle; since flow rate will be varying, it cannot be relied on as an indicator of when to reverse flow. A fixed cycle timer or the pressure drop must be used instead.

This modification will not apply to a system with a closed refrigerant loop, since (1) the deriming cycle probably will not be necessary in a properly designed closed loop system, and (2) the refrigerant flow rate will probably be relatively constant in a closed loop system unless variable storage volume is provided.

2.1.1 Note on Compressor Control

Unless a compressor is able to shed load and back out reduced air flow, reducing compressed air consumption cannot produce savings. Otherwise, the air will still be compressed even though it is not needed by the oxygen system. The best way to realize savings is by using a variable speed driver (motor-inverter, etc.) to match power supplied to compressor load. A fixed speed rotary compressor must be specially equipped with adjustable inlet guide vanes or for suction throttling if its capacity is to be adjusted. A reciprocating compressor can be equipped with automatic valve unloaders to provide capacity control. A less efficient alternative is to use a large supply tank to smooth supply pressure and run a fixed speed compressor in an on/off mode to regulate the air supply.

Aircraft systems using bleed air may be more likely to benefit from reducing air consumption, since air flow reductions translate directly to savings.

2.2 Refrigerant Flow Control (for Coolant/Oxygen Ratio)

When little oxygen is being condensed, coolant flow can be reduced. If the coolant compressor can shed the load reduction, this should reduce the energy consumed by the compressor.

A simple ratio will not necessarily maintain the exchanger mass at a temperature cold enough to resume operation without supplemental cooldown time. Coupling temperature control (based on TC1 and TC2 coolant outlet temperatures) with the flow ratio controls should keep the exchangers cold and permit rapid resumption of liquefaction (Figure 3). The control system will need to choose between the output of the flow ratio controller and the temperature controller; using the larger signal will insure minimum flow is maintained.

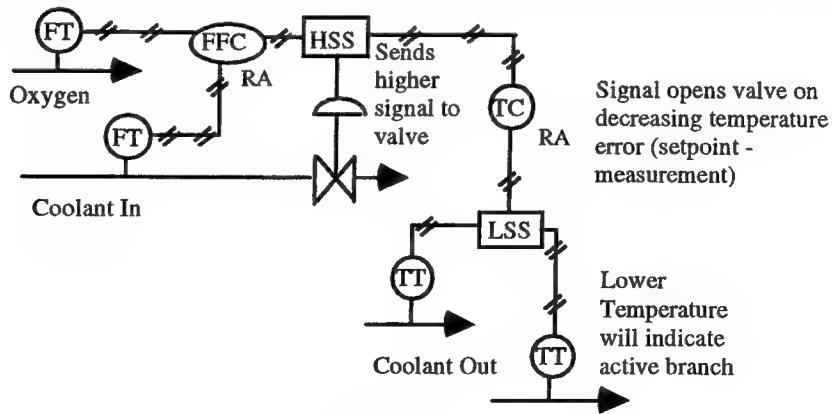


Figure 3

2.3 Concentrator Pressure Control

Pressure control of the concentrator inlet (near ASV1) might be useful, since it would permit decoupling of the concentrator, coolant, and (in future) instrument air supply pressures and allow optimization of the concentrator inlet pressure (a performance variable [9]). The present piping requires the concentrator and cooling air to run at the same pressure. Pressure control in the direction opposite to flow would probably be the best choice with respect to inventory control system consistency (Figure 4). The pressure control valve cannot replace the existing valve (ASV1), since valves designed for throttling and control are not capable of tight shutoff.

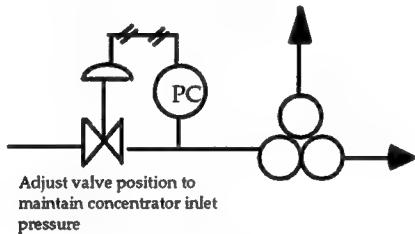


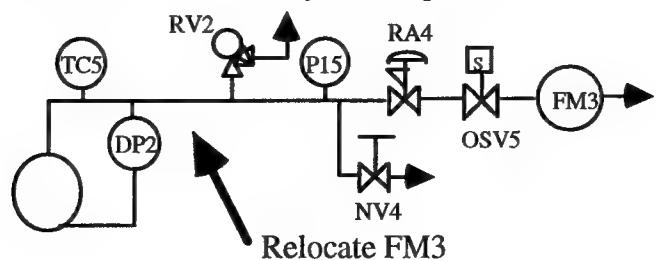
Figure 4

Selection of the concentrator operating pressure depends on more than the concentrator specifications. If the concentrator is to be bypassed directly to the demand users (Section 2.6), its pressure must be kept high enough to be compatible with the breathing regulators.

2.4 Oxygen Vent Control

It might be worthwhile to arrange to measure the amount of gas vented through the pressure relief valves (RV1, RV2) and the oxygen vent line (OSV3). Otherwise, it may prove difficult to develop good mass balances around the dewars. This would require three additional flow meters.

Alternatively, existing flow meters (FM3, FM4) could be relocated upstream of the relief valves (Figure 5). This would not distinguish between dewar vent and relief flows when the vent was open, but would permit measurement of the relief flow when the vent was closed. Similar relocation of the draw flow measurement device (FC1) is probably not wise, since it would interfere with the ability to manipulate the demand flow.



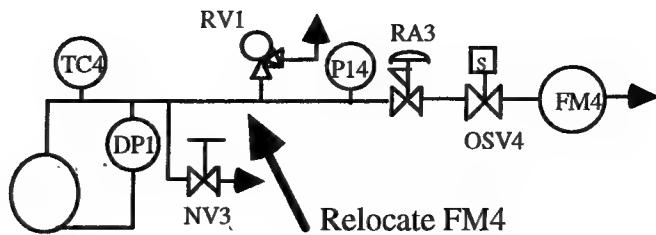


Figure 5

If it becomes apparent that the oxygen vent line (OSV3) is going to be used to bleed oxygen for control purposes, rather than strictly for emergency conditions, it should be replaced with a throttling valve.

The possibility of recovering vented oxygen is addressed in Section 2.9.

2.5 Oxygen Feed Rate & Concentrator Product Composition

The current operating plan is to open the oxygen supply valve (OSV1) as a step input [6]. The resulting flow increase will probably cause a temporary decline in oxygen concentration. If the drop is substantial or sustained, it might be wise to ramp up the oxygen flow on startup. A sufficiently slow ramp should keep the concentration closer to the desired level. The existing flow controller (FC2) should be adequate for this service. Tests on the Essex concentrator installed in HOS indicate that it responds fairly rapidly to flow rate changes; consequently, ramping flow on startup is probably unnecessary.

Oxygen flow may also be varied to maintain concentrator outlet composition during a run. This is probably best achieved by cascade connection of an analyzer controller to the existing flow controller (Figure 6). Suitable overrides to close the oxygen valve when the dewars are full will also be needed.

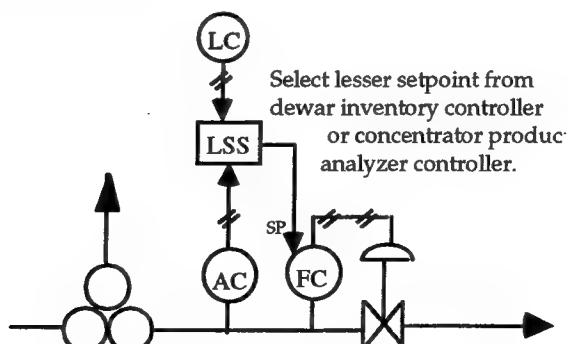


Figure 6

If the oxygen concentration is controlled in this manner, a separate ramping function for concentrator startup won't be needed; the controller should be adequate.

2.6 Bypass of Concentrator Product to Users

It might be desirable to bypass the liquefier

- 1) in case of mismatch between LOX converter production and demand
 - to meet temporary demand overloads
 - to provide oxygen immediately on compressor startup and before the liquefier has had time to cool down.
- 2) for LOX conservation
 - to conserve inventory for scheduled downtime; for example, to provide a pilot with breathing gas prior to engine startup
 - to maximize inventory available for an anticipated critical use interval or as insurance against emergency demands
- 3) to operate the concentrator in an off-design mode for air purification rather than oxygen concentration

All these cases depend on reconciling the operating pressures of the concentrator product, dewar/converter product, and the demand.

The optimum operating range for most concentrators seems to be in the 30-50 psig range. HOS is designed to supply oxygen from its dewars at about 70 psia, to insure compatibility with standard breathing regulators [2]. Higher pressures may be advantageous, since higher temperatures may then be used. The design point for the compressed air supply to HOS is 55 psig [7].

Any bypass arrangement must coordinate these operating pressures so that the demand can be met. This may require raising the operating pressure of the concentrator, reducing the operating pressure of the dewars, or using lower pressure breathing regulators to reduce the demand pressure. Once these issues are resolved, it should be straightforward to devise logic to switch between direct converter flow and concentrator product.

The most straightforward approach to developing such a bypass requires operating so that the concentrator outlet pressure is greater than or equal to the demand pressure. Bypass flow can then pass through a pressure control valve to merge with converter oxygen as it is released by the dewar pressure regulator (Figure 7).

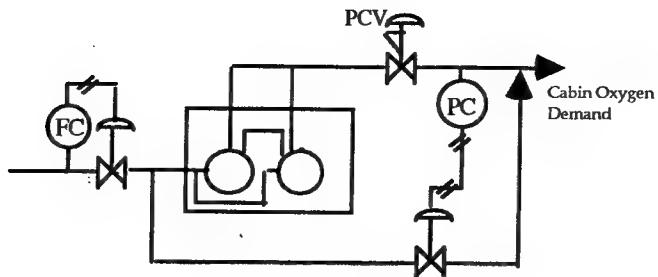


Figure 7

When demand exceeds converter production, the controller will open the bypass; if converter production is adequate or high, the pressure will rise and the controller will close the bypass valve. Such a pressure controller should take care of situations where the concentrator was running but the liquefier was not able to meet demand. This system would require the bypass branch off after the concentrator flow controller, so that it could be used to manipulate oxygen concentration.

Operation in off design mode could be achieved by blocking out the liquefier inlet; with no LOX being produced by the converters, the pressure controller would open the bypass fully. The concentrator flow controller would need to be reset to match the changed operating conditions.

LOX conservation is more complicated, and depends mostly on the LOX converter design and operating procedure. Conservation requires removing oxygen from the converters just fast enough to prevent overpressure or at a rate well below the normal vaporization rate, to maintain the largest inventory of liquid. If variable input heaters are used to drive vaporization, they will need to be operated near minimum. When converter make is reduced, the demand pressure controller would compensate by increasing bypass flow.

2.6.1 LOX Storage Concerns

One set of operating scenarios for the HOS systems calls for storage of LOX beyond what is needed to meet the immediate demand. The current configuration does not seem to address this issue, so it seems wise to run some test cases to determine how best to modify the system and adjust planned operation.

During storage, LOX will vaporize due to heat transfer to the dewar from outside. This will lead to vaporization and venting of oxygen for pressure regulation. The ADL report provides information that can be used to get a crude estimate of how long the dewars can

retain oxygen. ADL calculated 40 W of residual heat input contributing to dewar self-pressurization [2]. If all heat is transferred to the dewars and used to vaporize oxygen,

$$\left(40 \frac{\text{J}}{\text{s}}\right) \left(\frac{\text{g}}{199.7 \text{ J}}\right) = 0.20 \frac{\text{g}}{\text{s}}$$
 is vaporized. To maintain constant dewar pressure, this additional vaporization must be vented. If the density of LOX is 1100 g/l, then vaporization reduces the stored LOX by 0.66 l/h. Consequently, a full 5 liter dewar will last for only 7.6 hours at HOS conditions. The 40 W number is based on continuing operation of the refrigeration system, so actual heat losses might be much higher.

If variable pressure operation of the oxygen side has been provided (Section 2.8), it might make sense to allow the dewars to pressurize to a "storage pressure" (say 7 atm) greater than the normal draw pressure (5 atm) whenever it is desired to retain LOX, reverting to the draw pressure when the dewars are used to supply the cabin demand. Available LOX converters appear to be able to operate up to 120 psig (~9 atm absolute), so the higher pressure seems feasible.

2.7 Dewar Switching for Continuous Flow

In a medical application, it may be critical to maintain a continuous flow of product oxygen. Consequently, the dewar switching sequence will need to be examined. The initial plan is to fill one dewar while drawing from the other and to switch both dewars at once [6]. Once a dewar is switched to the higher draw pressure, it requires time to self-pressurize to that pressure. During this interval, the product flow will be reduced or eliminated. Consequently, for some interval after the switch, flow will be disrupted. HOS test runs should be conducted to determine whether this pressurization interval is significant.

If pressurization time is of concern, the dewar switching pattern can be modified to support continuous flow. As the draw dewar (1) nears empty, the other dewar (2) should be switched to the draw mode and the oxygen flow to the liquefier switched off. Both dewars will then be operating at the 5 atm draw pressure. Once the switched dewar (2) has fully self-pressurized, so that it is capable of supplying the full demand, the original dewar (1) can be switched to the fill mode and the oxygen flow restored. Thus, by overlapping the draw phases of the dewars, flow continuity can be maintained.

It will be interesting to examine the fill and draw intervals produced by HOS during operation. A 5 liter dewar should hold about 5.5 kg of LOX (using 1100 kg/cubic meter as the LOX density). At the ADL nominal design filling rate of 0.5 g/s [2, p. 46], 183 minutes are needed to produce this mass (224 min. at the ADL Run 7 average rate of 0.41 g/s), plus whatever additional filling is required to cool the dewar. At 1 atm and 273 K, 5.5 kg of oxygen occupies 3852 liters (assume ideal gas). If this is drawn off at 30 l/min., it will last for 128 minutes. The 55 minute difference between the fill interval and the draw interval could make it difficult to sustain operation for significant time periods (greater than 4 hours). The draw rate would have to be reduced to 21 l/min. to balance the fill and draw intervals.

2.8 Variable Oxygen Side Pressure

The upper limit on the liquefier operating pressure is set by the pressure required to operate a high pressure breathing regulator. ADL refers to regulators that require 70 psi inlet gas [2].

The lower limit is less firmly established. Cooling limits impose a minimum, but if 90 K can be reached, atmospheric operation is not impossible. ADL data suggests that operation below 2 atm may be cooling limited [2]. ADL set the vent line regulators at 2 atm; operating data shows that the dewar pressure generally ranged from 5 to 20 psig, although one run briefly reached 30 psig. The current HOS operating plan is to fill dewars at 3 atm [4,6].

It would probably be worthwhile to try to determine whether an optimum pressure sequence exists. Compare LOX make, air consumption, cycle times, etc., to quantify the costs of lower pressure operation. The HOS test program ought to look at operating at a range of pressures (covering at least 2-4 atm) to determine the acceptable operating range. Particular attention should be paid to the minimum workable pressure difference between the fill and draw modes. The current plan provides a difference of 2-3 atm; a smaller difference would probably increase operational flexibility.

A better understanding of the relationship between oxygen side pressure and performance will be useful when addressing problems related to bypassing the liquefier (Section 2.6) and switching of dewars (Section 2.7). Variable pressure operation of the oxygen side might be useful in efforts to match production and demand rates. Optimization of the

operating pressure is possible with the current manual setup, or dewar vent line regulators RA3 and RA4 could be replaced with pressure controlled valves to allow the set point pressure to be varied more easily.

2.9 Vent Gas Recovery

Reducing or recovering oxygen lost through the dewar vents can produce savings. Each liter of oxygen entering the dewars requires that air be compressed and run through the concentrator. If the concentrator were perfectly efficient, stoichiometry would require that 4.5 liters of air be processed for every liter of 95% oxygen produced. Reducing venting serves to recover the compression energy as well as the energy used to liquefy the oxygen.

The dewars vent gaseous oxygen periodically during operation. When filling, the dewars are cooled by vaporizing LOX that then must be vented. A fully pressurized converter also vents some oxygen to maintain pressure and keep lines open. The energy losses from the vent might be partially offset by recovering the vented oxygen and returning it to the concentrator outlet. Except on initial startup, the vented material will be oxygen, so recycling the gas seems feasible. Since the vent line pressure will swing, check valves will be necessary (Figure 8). Vent gas recycle will reduce the load on the air compressor from the concentrator and should also modestly reduce the amount of cooling required for liquefaction since the recycle will probably be cooler than the concentrator product.

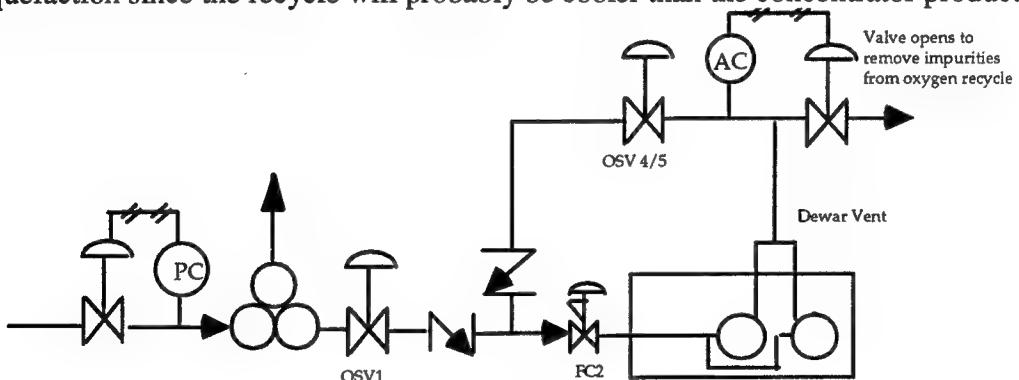


Figure 8

The main obstacle to vent recycle is balancing the system operating pressures. Dewar filling pressure (planned to be 3 atm [6]), must be high enough to enter the concentrator product stream, which may operate as high as 60 psig under the current plan; thus to implement vent recycle, concentrator pressure control (Section 2.3) would probably be essential. Also note that the pressure requirements of vent recovery might conflict with

those needed to bypass the liquefaction section (Section 2.6). The recycle loop will need to have a purge connection to eliminate air and contaminants from the stream. This can be done periodically, or an analyzer based controller can perform the function automatically. Additional over pressure protection can be added in the form of a high pressure override on the purge valve.

Potential savings by vent recycle seem real. During ADL Run 7 [2], dewar vent flows averaged 74.4 g/min. for 138 minutes as venting ranged from 35.5 to 100.9 g/min. In the same run, product oxygen demand flow averaged only 50.2 g/min. for 72 min. It is worth noting, however, that ADL apparently never refilled a "cold" dewar, and so these data may not be a good representation of steady-state. HOS runs can be conducted to clarify the issue

2.10 Automated JT Valve

If complete automation is desired, the JT valve probably can be fitted with a rotary actuator and set from the computer; however, it may be hard to find an actuator that will work with the existing valve. Although the JT valve is not part of the modified HOS plan, there are no current plans for its removal. Since impurities will tend to cause sticking of the JT valve, any actuator would need to be direct acting or employ a positioner. Most likely, an actuator cannot be justified.

3.0 Conclusions

A number of opportunities exist to improve operation of the hybrid oxygen systems. Small modifications in piping or control systems should produce appreciable decreases in compressed air consumption and consequently improve energy use efficiency.

Psychological Issues in Pilot Training Completion and Retention

Paul D. Retzlaff
Associate Professor
Department of Psychology

University of Northern Colorado
Greeley, CO 80639

Final Report for:
Summer Faculty Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base
Washington, D.C.

and

Armstrong Laboratory

October 1995

Psychological issues in pilot training completion and retention.

Paul D. Retzlaff
Associate Professor
Department of Psychology
University of Northern Colorado

A number of studies have examined the intelligence and personality of pilots. Few, however, have been able to utilize long-term follow-up data. Three-hundred and fifty Air Force officers undergoing Undergraduate Pilot Training were administered the Multidimensional Aptitude Battery, the Personality Research Form, and the Millon Clinical Multiaxial Inventory. Ten year follow-up data is provided on pilot training completion and length of service. No differences were found among the training completions groups but a number of consistent personality variables were correlated with length of service.

Psychological issues in pilot training completion and retention.

Paul Retzlaff

Reviews of the literature on personality and pilot personnel issues have generally found few predictive relationships (Dolgin and Gibb, 1988). No single test or variable has emerged which is widely accepted as adding a great deal of predictive validity to pilot selection or training outcome. Still, there remain a few studies which suggest a number of global personality characteristics such as self-confidence and flexibility that can aid, to a small degree, in the prediction of such things as training outcome in the US Air Force (Siem, 1992). It is difficult to know if the lack of personality prediction is a function of true lack of association, poor psychological tests, or limited and early outcome data. No study thus far has looked at multiyear follow-up.

Retzlaff and Gibertini (1988) presented psychological data on 350 US Air Force pilot training students. Testing included scales of intelligence, personality, and psychopathology. The Multidimensional Aptitude Battery (MAB; Jackson, 1985) was used to assess intelligence. Full Scale IQ was found to be 120 for the group. The ten subtests showed mean performances about one standard deviation above the normative mean. The Personality Research Form (PRF; Jackson, 1984) provided data on personality characteristics. Student pilots were found to be higher on scales of affiliation, cognitive structure, dominance, and social desirability than a college student control sample. The student pilots were also lower on scales of abasement, autonomy, harm avoidance, and understanding than the other sample. Finally, the Millon Clinical Multiaxial Inventory (MCMI; Millon, 1983) was used to assess psychopathology. The only findings here was

evidence of histrionic and narcissistic personality features. All of these data were provided as a norm sample against which clinically referred pilots might be compared.

Additional statistical work pointed to three prototypical types of aviator (Retzlaff and Gibertini, 1987). The first cluster was dubbed the "Right Stuff". This group was aggressive, dominant, exhibitionistic, impulsive, and playful. The second group, the "Company-man", had more achievement, endurance, affiliation, and order traits. Finally, the last group, the "Wrong Stuff", had the lowest scores of all on exhibition, understanding, affiliation, and achievement. This work was cross-validated by method, sample, and test showing a remarkable stability of structure. While making intuitive sense, no follow-up data was available to these researchers at that time to validate this typology.

Purpose

The purpose of the current study was to follow-up the 350 student pilots tested by Retzlaff and Gibertini. Specifically, completion of Undergraduate Pilot Training (UPT), subsequent aviation rating, and 10 year service retention were of interest.

METHOD

Subjects

Subjects for this study were 350 white males who entered United States Air Force Undergraduate Pilot Training at Reese Air Force Base, Texas between December 1984 and September 1985. Ages at the time ranged from 22 to 27 years. Subjects were administered the Multidimensional Aptitude Battery (N=145), Personality Research Form (N=340), and Millon Clinical Multiaxial Inventory (N=249) in classes

of 50 to 55 within the first 4 weeks of training.

Approximately ten years after testing (June 1995), UPT completion, aviation rating, and length of service were extracted from computerized records of active duty officers and separated officers by the Armstrong Laboratory, Decision Support Branch technical staff. UPT completion was indicated by having an aviation rating of pilot. Rated navigators either were so rated prior to UPT and failed pilot training or went to Undergraduate Navigator Training after an unsuccessful attempt to complete UPT. Those with no rating were found in a wide range of AFSC's after their UPT experience.

Measures

The Multidimensional Aptitude Battery (MAB; Jackson, 1985) is a broad based test of intellectual ability. It was patterned on the Wechsler Adult Intelligence Scale-Revised (WAIS-R), the most widely used individually administered test of intelligence. While the WAIS-R requires about an hour and a half per subject to individually administer, the MAB can be given to large groups in about the same amount of time. Additionally, the WAIS-R requires skillful scoring while the MAB has a multiple choice format. All subtests in the WAIS-R have corresponding paper and pencil subtests in the MAB except immediate digit memory. Verbal components tapped include information (INFO), comprehension (COMP), arithmetic (ARITH), similarities (SIM), and vocabulary (VOCAB). Performance measures include digit symbol coding (DIGSYM), picture completion (PIXCOMP), spatial (SPAT), picture arrangement (PIXARR), and object assembly (OBJASS). Scores on each of the subtests are scaled to a mean of 50 and a standard deviation of 10. Verbal (VERBAL) and performance (PERF) IQ's are available as is a full scale (FULL) IQ, each scaled to a mean of 100 and a standard

deviation of 15. Reliabilities for the summary scores range from .94 to .98, and the correlation with the WAIS-R is .91.

The Personality Research Form (Jackson, 1984), Form E, is a 352 true-false item, 21 scale inventory of normal personality traits. Twenty scales tap domains of interest in the aerospace environment including abasement (AB), achievement (AC), affiliation (AF), aggression (AG), autonomy (AU), change (CH), cognitive structure (CS), defencence (DE), dominance (DO), endurance (EN), exhibition (EX), harm avoidance (HA), impulsivity (IM), nurturance (NU), order (ORD), play (PL), sentience (SE), social recognition (SR), succorance (SU), and understanding (UN). The twenty-first scale is a validity scale which attempts to assess highly infrequent (IN) item response indicative of random test response. The twenty-second scale assesses the degree of social desirability (SD) of the subject's responses. Scores range from 0 to 16 on each of the scales. Reliabilities range from .57 to .91.

The Millon Clinical Multiaxial Inventory (MCMI) (Millon, 1983) is a 20 scale instrument which was developed to measure dimensions consistent with modern psychiatric diagnostic nomenclature (DSM-III; American Psychiatric Association, 1980). It assesses eight basic personality patterns (schizoid (SC), avoidant (AV), dependent (DD), histrionic (HI), narcissistic (NA), antisocial (AN), compulsive (CO), and passive aggressive (PA)), three pathological personality patterns (paranoid (PA), schizotypal (ST), and borderline (BO)), and nine clinical syndromes (anxiety (AX), somatoform (SO), hypomania (HY), dysthymia (DY), alcohol use (AL), drug use (DR), psychotic thought (PT), psychotic depression (PD), and delusions (DL)). The last scale

is a weight (WE) indicator of over and under reporting psychopathology. The cut-off score for clinically significant elevations is 74 on all scales. Scores above 84 indicate the most prominent elements in the patient's clinical picture. While this instrument was developed for use with pathological populations, its basic personality scales are useful in normal groups (Retzlaff and Gibertini, 1987) and its clinical syndrome scales may provide a good screening for mood disruptions such as anxiety and depression. It must be noted that the alcohol, drug, and psychotic scales of this version of the MCMI are extremely poor psychometrically (Gibertini, Brandenburg, and Retzlaff, 1986) and should not be used with student pilots or other high functioning individuals. Basic personality scale reliabilities range from .73 to .91.

RESULTS

Table 1 presents the descriptive statistics for the MAB for not only the entire sample but also for the three training outcome groups individually. As can be seen, the Full Scale IQ scores of the entire sample are well above average at 120. Performance scores are slightly above Verbal IQ scores. The subtest scores are T statistics with a mean of 50 and a standard deviation of 10 in the normative sample. Additionally, all pilot subtest scores are about 1 standard deviation above the normative sample.

The only significant difference among the three outcome groups is on the Arithmetic variable with those becoming pilots having the highest scores. The three point difference between the highest and lowest group, however, is quite modest. There are no differences in any of the other summary or subtest scores across the three groups. Those who became pilots, navigators, or who were not subsequently

rated showed few differences on this cognitive testing.

PRF personality variables are presented in Table 2. Again, these student pilots are higher than college students on affiliation, cognitive structure, dominance, and social desirability. They are lower on abasement, autonomy, harm avoidance, and understanding (Jackson, 1984; Retzlaff and Gibertini, 1987). This finding, of course, may be a reflection of differences in age, education, or other moderating variable between student pilots and college students.

The mean raw scores for the three outcome samples are not statistically different. Interestingly, the largest, though nonsignificant, difference is on the Achievement variable.

The MCMI (Table 3) points to predominant histrionic and narcissistic patterns of personality. This personality constellation is highly consistent with lay impression of the pilot as a highly sociable person who has strong self-esteem. Very little, if any, clinical pathology is seen. Although base rate scores in the clinical range (BR > 84) were found on 9 scales, only 3 of these (Narcissistic, Histrionic, and Antisocial) had positive rates greater than the rate of false positives (Millon, 1983) found in clinical samples. While "psychiatric pathology" is not expected, moderate elevations are usefully interpreted. Indeed, some scales' elevation or lack of elevation may be more meaningful than others. King (1994) points to typical aviator elevations on Narcissistic, Histrionic, Compulsive, and Antisocial. And as such, elevations on scales such as Schizoid or Avoidant would be of far greater clinical concern.

Again, no training outcome differences are seen across the three groups. The "Right Stuff" variables do not seem to be differentiating

the groups as to "getting their wings".

Table 4 provides the data on the retention outcome variable as a function of the MAB. Here length of service was trichotomized into three service commitment blocks. The first block of 5 and less years of service is indicative of initial separation from service as well as those not completing flight training but remaining on in the service briefly in some other occupational capacity. The second block of 6 to 9 years is typical of those completing UPT, "earning their wings", and only serving out their initial commitment. The final block is reflective of those who chose to stay on in the AF beyond their initial commitment. Almost invariably these were pilots as will be seen later.

Comprehension showed a statistically significant difference among groups. This is probably due to a statistical artifact. There were only 13 subjects in the less than 5 years group and they all had very similar Comprehension scores. This similarity resulted in a standard deviation of only 1 point. This no doubt lead to the statistical significance. No summary IQ scores or other scale scores rose to the level of statistical significance across the three groups.

Table 5 provides the total and three subsamples' PRF means and standard deviations. Here Play showed significant differences across the three groups. The least Playful are found to remain on active duty 10 years after training. The most Playful stayed in the service the least amount of time. Other variables which approached significance were Achievement and Social Desirability. The "still on AD" group had the highest Achievement scores and Social Desirability scores.

Table 6 presents the length of service outcome by MCMI. Only one

significant difference was found across the three groups. The group spending the least amount of time in the service was the most Dysthymic. These scores, however, are well below the clinically relevant range and therefore may reflect some subclinical narrow aspect of dysthymia. As testing was accomplished well before any training began, this is probably not a situational reaction to imminent discharge. The largest, although non-significant, difference is found on the Compulsive variable with those highest spending the most time in service. This would be consistent with the Play results of the PRF.

DISCUSSION

The most dramatic finding of these outcomes measures is how little of pilot training completion and service retention is predicted by the intelligence, personality, and psychopathology testing instruments. Large amounts of variance are not predicted.

The pilot training outcome data appears to be generally independent of the intelligence testing given to the officers as students. Some differences may be related to arithmetic ability, although this must be replicated.

Indeed, the current results are not a reason to believe cognitive variables such as IQ are not important in the completion of pilot training. The IQ's of these subjects are well above the population as a whole. It in many ways is a statistical forgone conclusion that IQ will not predict training outcome because many of the screening criteria for pilot training opportunity are highly correlated with IQ. For example, admission to the Air Force Academy is contingent upon high school grades and college entrance tests, both highly correlated

with IQ. Selection for pilot training after the Academy is based upon the technical nature of the undergraduate major and undergraduate grades. Again, both variables are highly correlated with IQ. In essence, IQ is probably so predictive of pilot training outcome that it is already a major, albeit embedded, portion of the selection criteria.

Neither the test of personality nor the test of psychopathology predicted training outcome. It must be pointed out, however, that all students failing pilot training for any reason were combined. This was done in order to have a sufficiently large sample of training failures to meet the statistical design requirements. In doing so, however, those who were eliminated for poor flying, which was the vast majority, were combined with a number of students who were eliminated due to medical conditions (e.g., high blood pressure), fear of flying, air sickness, or administrative problems. It is apparent that a very large initial sample is necessary to separate out all of these reasons for UPT noncompletion.

More interesting data is seen in the service retention numbers. Intelligence testing is of limited outcome value for many of the same reasons stated above. Here small, but consistent, findings are seen across personality tests and across methodologies. It would appear that playful, short-sighted individuals are not retained in the AF. Here, achievement-oriented, conforming, and compulsive-type individuals tend to remain on active duty past their initial "pay back" commitments. With the cost of training new pilots and the consequent loss of experience, retention potential may be of particular interest.

REFERENCES

American Psychiatric Association. (1980). Diagnostic and statistical manual of mental Disorders (3rd ed.). Washington, D.C.: Author.

Costa, P. T. & McCrae, R. R. (1992). NEO PI-R: Professional Manual. Odessa, FL: Psychological Assessment Resources.

Dolgin, D. and Gibb, G. D. (1988). Personnel assessment and aviation selection: Past, present, and future. In R. S. Jensen (Ed.), Aviation Psychology (3rd ed., pp. 288-320). London: Gower.

Flynn, C. F., Sipes, W. E., Grosenbach, M. J., and Ellsworth, J. (1994). Top performer survey: Computerized psychological assessment of aircrew. Aviation, Space, and Environmental Medicine, 65, 39-44.

Gibertini, M., Brandenburg, N., and Retzlaff, P. D. (1986). The operating characteristics of the Millon Clinical Multiaxial Inventory. Journal of Personality Assessment, 50, 554-567.

Jackson, D. N. (1984). Personality Research Form. Port Huron, MI: Research Psychologists Press.

Jackson, D. N. (1985). Multidimensional Aptitude Battery. Port Huron, MI: Research Psychologists Press.

King, R. E. (1994). Assessing aviators for personality pathology with the Millon Clinical Multiaxial Inventory (MCMI). Aviation, Space, and Environmental Medicine, 65, 227-231.

King, R. E. and Flynn, C. F. (in press). Defining and measuring the "Right Stuff": Neuropsychiatrically Enhanced Flight Screening (N-EFS). Aviation, Space, and Environmental Medicine.

Millon, T. (1983). Millon Clinical Multiaxial Inventory. Minneapolis, MN: National Computer Systems.

Retzlaff, P. and Gibertini, M. (1988). The objective psychological testing of Air Force officers in pilot training. Aviation, Space, and Environmental Medicine, 59, 661-663.

Retzlaff, P. and Gibertini M. (1987). Air Force pilot personality: Hard data on "The Right Stuff". Multivariate Behavioral Research, 22, 383-399.

Siem, F. M. (1992). Predictive validity of an automated personality inventory for Air Force pilot selection. The International Journal of Aviation Psychology, 2, 261-270.

Table 1

MAB Means, SDs, and F tests for UPT completion.

	ALL		Pilots		Navigators		No Rating		F(2,130)
	Mean	sd	Mean	sd	Mean	sd	Mean	sd	
INFO	57	6	57	5	60	4	57	6	1.76
COMP	57	4	57	4	56	8	58	2	0.31
ARITH	62	5	62	5	60	4	59	7	3.44*
SIM	58	6	58	5	60	4	56	9	1.84
VOCAB	57	6	57	7	59	4	55	6	1.46
DIGSYM	61	7	61	7	61	6	61	8	0.05
PIXCOMP	57	5	58	5	58	4	56	6	0.72
SPAT	61	7	61	7	61	7	60	5	0.15
PIXARR	63	8	63	8	62	10	64	7	0.29
OBJASS	58	6	58	6	56	6	59	5	0.69
VERBAL	116	7	116	7	119	6	114	8	1.44
PERF	121	10	121	10	120	12	121	10	0.06
FULL	120	8	120	8	121	9	119	9	0.30
N	133		94		14		25		

Note: * .05 level of significance = 3.07, .01 = 4.78.

Table 2

PRF Means, SDs, and F tests for UPT completion.

	ALL		Pilots		Navigators		No Rating		F(2,298)
	Mean	sd	Mean	sd	Mean	sd	Mean	sd	
AB	6.1	2.5	6.3	2.4	5.7	2.8	5.7	2.8	1.54
AC	12.0	2.6	12.2	2.5	11.8	3.1	11.3	2.8	2.87
AF	10.9	3.3	10.9	3.2	11.5	3.5	10.4	3.3	0.93
AG	7.5	3.0	7.5	2.9	7.7	3.0	7.6	3.3	0.10
AU	6.7	2.9	6.6	2.9	7.2	3.3	6.7	2.7	0.39
CH	9.5	2.6	9.6	2.6	10.5	1.9	9.0	2.8	2.85
CS	10.2	2.8	10.2	2.7	10.3	3.1	10.3	2.9	0.05
DE	5.5	3.2	5.6	3.3	5.0	2.4	5.6	3.1	0.41
DO	12.6	2.5	12.5	2.6	13.4	1.9	12.8	2.4	1.39
EN	11.2	2.5	11.3	2.4	10.3	2.9	11.1	2.5	1.41
EX	9.1	3.8	9.2	3.8	9.6	3.2	8.6	4.1	0.74
HA	5.5	3.0	5.5	2.8	5.5	3.0	5.5	3.5	0.01
IM	5.0	3.4	5.1	3.1	4.7	4.2	4.9	3.8	0.13
NU	9.6	2.7	9.7	2.8	9.1	3.0	9.5	2.2	0.59
ORD	8.4	4.0	8.3	4.0	8.7	4.6	8.7	4.0	0.24
PL	8.6	2.9	8.7	2.9	7.8	2.4	8.6	3.4	0.88
SE	8.8	2.9	8.7	3.0	8.4	2.7	9.4	2.6	1.66
SR	8.8	3.0	8.9	3.0	9.1	3.3	8.3	3.2	0.80
SU	5.8	3.4	5.8	3.4	5.7	2.5	5.8	3.8	0.01
UN	8.1	3.1	7.9	3.0	8.6	3.1	8.6	3.2	1.29
IN	0.3	0.7	0.2	0.6	0.1	0.3	0.4	0.9	1.60
SD	13.5	1.9	13.5	1.9	13.2	2.2	13.3	2.0	0.64

N	301	219	23	59
---	-----	-----	----	----

Note: .05 level of significance = 3.03, .01 = 4.68. No F statistics are significant.

Table 3

MCMI Means, SDs, and F tests for UPT completion.

	ALL		Pilots		Navigators		No Rating		
	Mean	sd	Mean	sd	Mean	sd	Mean	sd	F(2,212)
SC	22	14	21	14	22	13	25	14	1.07
AV	19	16	18	16	16	15	22	17	1.30
DD	42	19	43	19	38	15	41	22	0.45
HI	71	17	71	17	69	20	71	15	0.07
NA	72	15	72	15	70	13	71	14	0.13
AN	65	15	65	15	64	11	65	15	0.03
CO	67	10	68	10	65	13	66	10	0.73
PA	24	16	24	15	21	19	26	18	0.78
ST	31	18	30	18	31	19	33	18	0.35
BO	35	17	34	17	33	16	37	18	0.68
PR	59	13	59	13	57	9	60	13	0.21
AX	44	21	44	21	40	20	48	19	1.09
SO	54	15	54	16	51	17	55	14	0.51
HY	43	27	45	27	32	26	42	28	1.65
DY	42	20	40	20	43	17	48	19	2.28
AL	32	14	32	14	26	13	33	14	1.71
DR	54	17	54	18	52	17	55	16	0.24
PT	32	19	32	19	29	20	35	20	0.53
PD	23	19	21	19	21	19	28	18	2.20
DL	48	18	49	18	41	17	48	16	1.39
WE	1	2	1	2	1	1	1	2	0.04

N	215	151	17	47
---	-----	-----	----	----

Note: .05 level of significance = 3.04, .01 = 4.71. No F statistics are significant.

Table 4

MAB Means, SDs, and F tests for Length of Service

	All		0 to 5		6 to 9		Still AD		
	Mean	sd	Mean	sd	Mean	sd	Mean	sd	F(2,132)
INFO	57	6	60	3	57	7	57	5	1.24
COMP	57	4	59	1	56	6	58	3	3.36*
ARITH	62	6	60	7	61	6	62	5	0.75
SIM	58	6	56	7	58	6	58	6	0.55
VOCAB	57	6	58	4	56	8	57	6	0.83
DIGSYM	61	7	63	7	59	8	61	7	1.48
PIXCOMP	57	5	58	4	57	5	58	5	0.17
SPAT	61	7	61	4	60	8	61	7	0.26
PIXARR	63	8	66	6	62	8	64	8	1.28
OBJASS	58	6	59	5	57	7	59	5	1.18
VERBAL	116	7	117	5	115	8	117	7	1.45
PERF	121	10	123	9	118	12	121	9	1.63
FULL	120	8	122	7	117	9	121	8	2.32
N	135		13		43		79		

Note: * .05 level of significance = 3.07, .01 = 4.78.

Table 5

PRF Means, SDs, and F tests for Length of Service

	All		0 to 5		6 to 9		Still AD			F(2,301)
	Mean	sd	Mean	sd	Mean	sd	Mean	sd	F(2,301)	
AB	6.1	2.6	6.1	3.0	6.2	2.5	6.1	2.5	0.07	
AC	12.0	2.6	11.3	2.9	11.8	2.7	12.3	2.4	2.87	
AF	10.9	3.3	11.1	3.4	10.9	3.6	10.9	3.0	0.09	
AG	7.5	3.0	6.9	3.4	8.0	3.0	7.4	2.9	2.09	
AU	6.6	2.9	7.0	3.1	7.0	2.9	6.3	2.8	1.94	
CH	9.5	2.6	9.2	2.7	9.6	2.6	9.5	2.5	0.46	
CS	10.2	2.8	10.2	2.6	9.9	3.1	10.5	2.6	1.41	
DE	5.5	3.2	4.6	2.5	5.8	3.6	5.6	3.0	1.92	
DO	12.6	2.5	12.5	2.9	12.7	2.4	12.6	2.5	0.13	
EN	11.2	2.5	11.0	2.6	11.0	2.3	11.4	2.5	0.87	
EX	9.1	3.8	9.6	4.5	9.5	3.8	8.8	3.6	1.40	
HA	5.5	3.0	5.6	3.0	5.2	2.9	5.7	3.0	0.90	
IM	5.0	3.4	5.2	3.4	5.4	3.6	4.7	3.2	1.56	
NU	9.6	2.7	9.2	2.4	9.6	2.8	9.8	2.7	0.78	
ORD	8.4	4.1	8.1	4.1	8.5	4.3	8.4	3.8	0.15	
PL	8.6	2.9	9.4	3.6	9.0	2.9	8.1	2.7	4.43*	
SE	8.8	2.9	9.3	2.6	9.0	3.0	8.6	2.8	1.25	
SR	8.8	3.0	8.0	3.1	8.8	3.1	9.0	2.9	1.64	
SU	5.8	3.4	5.7	3.7	6.1	3.1	5.7	3.5	0.38	
UN	8.1	3.1	8.5	3.3	7.9	3.2	8.2	2.9	0.61	
IN	0.3	0.7	0.3	1.0	0.3	0.7	0.2	0.5	0.09	
SD	13.5	1.9	13.5	1.9	13.1	2.2	13.7	1.7	2.52	

N	304	39	108	157
---	-----	----	-----	-----

note: * .05 level of significance = 3.03, .01 = 4.68

Subsequent stepwise discriminant function analysis results in only the inclusion of Play in the equation.

Table 6

MCMI Means, SDs, and F tests for Length of Service.

	All	0 to 5	6 to 9	Still AD					
	Mean	sd	Mean	sd	Mean	sd	F(2,213)		
SC	22	14	22	15	21	13	23	14	0.42
AV	19	16	20	17	19	17	19	16	0.10
DD	42	19	43	23	42	19	42	19	0.04
HI	71	17	76	17	71	18	69	16	2.36
NA	72	15	74	17	72	15	70	14	0.85
AN	65	15	65	18	65	14	65	14	0.03
CO	67	10	64	11	67	11	68	9	2.77
PA	24	16	28	19	24	16	23	14	1.24
ST	31	18	34	17	30	19	31	18	0.64
BO	35	17	37	17	36	16	33	18	0.96
PR	59	13	60	14	60	12	58	13	0.53
AX	44	21	46	19	45	22	44	21	0.23
SO	54	15	57	14	53	16	54	15	0.85
HY	44	27	48	32	44	26	42	26	0.57
DY	42	20	50	17	39	20	42	20	3.91*
AL	32	14	34	15	33	12	30	14	1.46
DR	54	17	58	16	55	17	53	18	1.22
PT	32	19	36	19	32	18	31	20	0.82
PD	23	19	28	18	23	19	21	19	1.72
DL	48	18	46	17	48	17	48	19	0.19
WE	1	2	1	2	0	1	1	2	0.67
N	216		37		80		99		

Note: * .05 level of significance = 3.04, .01 = 4.71.

Analysis of Loads on the Neck and Head Joints Due to G-Y Acceleration

Ali M. Sadegh
Professor
Department of Mechanical Engineering

The City College of
The City University of New York
New York, NY 10031

Final Report for:
Summer Faculty Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base
Washington, D.C.

and

Armstrong Laboratory

August 1995

ANALYSIS OF LOADS ON THE NECK AND HEAD JOINTS DUE TO G-Y ACCELERATION

Ali M. Sadegh
Professor and Chairman
Department of Mechanical Engineering
The City College of The City University of New York

Abstract

During an emergency ejection from aircraft, pilots are subjected to a high acceleration that could cause injuries. These injuries, particularly at the neck region, are exacerbated with the additional weight that is added to the head gear of the pilot such as night vision goggles and helmet-mount displays. There have been some studies on the head acceleration in the Z and X direction. The objectives of this study were a: to employ DYNAMAN and or ATB to validate the test results for the G-y acceleration, and b: to determine the loads and the torques at the neck-head and neck-upper torso joints. The experimental data were collected from the biodynamics responses of human volunteers during an acceleration in the Y direction of a sled at the sled track facility at Armstrong Laboratory at WPAFB. The sled was subjected to 4G to 7G and duration ranging from 31ms to 250ms in a total of 9 cells. Three segments model consists of head, neck and upper torso was employed. The experiment was validated by selecting the spring and the dashpot parameters of the ATB model so that the test results reasonably match with the experimental results. Using these parameters, the loads and the torques at the neck-pin and the head-pin joints were determined. Plots of the maximum loads and torques versus the amplitude and the duration of the acceleration of the sled were presented. Further analysis is needed to assess the cause of neck injuries and evaluation of human tolerance levels.

ANALYSIS OF LOADS ON THE NECK AND HEAD JOINTS DUE TO G-Y ACCELERATION

Ali M. Sadegh

INTRODUCTION

It is conceivable that a fighter pilot will be faced with an emergency situation that will require ejection from the cockpit. In the ejection process the pilot's torso is subjected to a high acceleration, on the order of 15 G's (representative of a B52 aircraft ejection seat (1)), see Figure 1. The time required for an ejection is on the order of milliseconds. The combination of the high acceleration and the extraordinarily short period of time applies extremely high magnitude of load to the spinal column particularly cervical spine. This load causes injuries; many of them occur in the lumbar and thoracic regions of the spine, and particularly the cervical spine and the neck(1).

In recent years pilots are flying with night vision goggles, and helmet-mounted displays that have significant advantages over panel-mounted instruments. These additional instrumentation have increased the mass of the helmet system which would increase the risk of a major neck injury in an emergency escape situation. Scientists and investigators at the Armstrong laboratory at Wright-Patterson Air Force Base have started a program to determine parameters for helmet design (2).

Recent advances in seat design and pilot training have added more restrain to the pilot's lumbar and thoracic regions thereby reducing the risk of lower back injuries. However, lack of restrain in the neck and head regions has made the pilot's cervical spin vulnerable to injuries. This is due to the fact that the pilot's head must have a reasonable freedom of motion in order to ensure an adequate field of view. In the ejection process, the initial orientation of the head and the direction of the acceleration vector plays an important role in the magnitude of the load that would be applied to the neck. That is to say that, the head acceleration varies considerably when the airplane is in roll or spin conditions.

There have been some studies and analyses of the head acceleration due to the acceleration in 'z

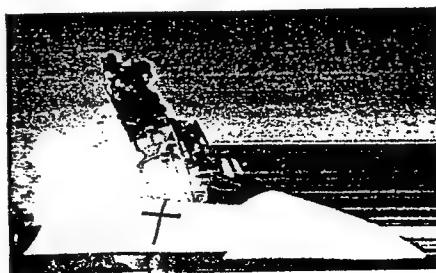


Fig. 1. Schematic of Pilot in Combination Ejecting from Cockpit

and x directions, (G-z and G-x), (3 and 4). However, during an acceleration in y direction (G-y), the acceleration and the movements of the head, and the load exerted to the neck, has not been fully studied. The purpose of this report is to validate the test results for the G-y acceleration by the Articulated Total Body (ATB) model, and also to determine the forces and the torques applied to the neck and head joints. The loads and the torques should be compared to human tolerance level .

OBJECTIVES:

The focus of this study was to employ the DYNAMAN or ATB as a tool for predicting the global rigid body dynamic response of human subjects in order to validate the result of the slid experiment in the y-direction, and to determine the loads and the torques exerted on the neck and the head . In particular:

- a: to validate and to select the spring and dashpot coefficients of the ATB model so that ATB's pattern of the accelerations of the head at the mouth pack match to the measured acceleration of the mouth piece of the test.
- b: to determine the loads and the torques at the neck/head and at the upper torso/neck joints. It is expected that the forces and the torques determined in this study will be used as a boundary condition of a more detailed modeling and analysis of the head/neck system.

EXPERIMENTAL DATA

At the sled track facility located at the Armstrong Laboratory at WPAFB, the experimental data were collected from biodynamic responses of human volunteers during an acceleration in the Y direction. This facility employs an Impulse Accelerator (Shaffer 1976) that is a gas powered actuator which accelerates a sled on a two-rail track. The volunteers were placed in a chair that is attached to a sled and facing perpendicular to the direction of the track. Two sets of 3 orthogonal linear accelerometers were located in a chest pack and a mouth pack. These accelerometers collected the X, Y and Z accelerations of the torso and head as a function of time during the acceleration impulse. The coordinate system of the acceleration and the chair is shown in Figure 2a. The sled was subjected to the acceleration pulse of a half-sine with peak acceleration ranging from 4G to 7G and duration ranging from 31 ms to 250 ms. Total of 9

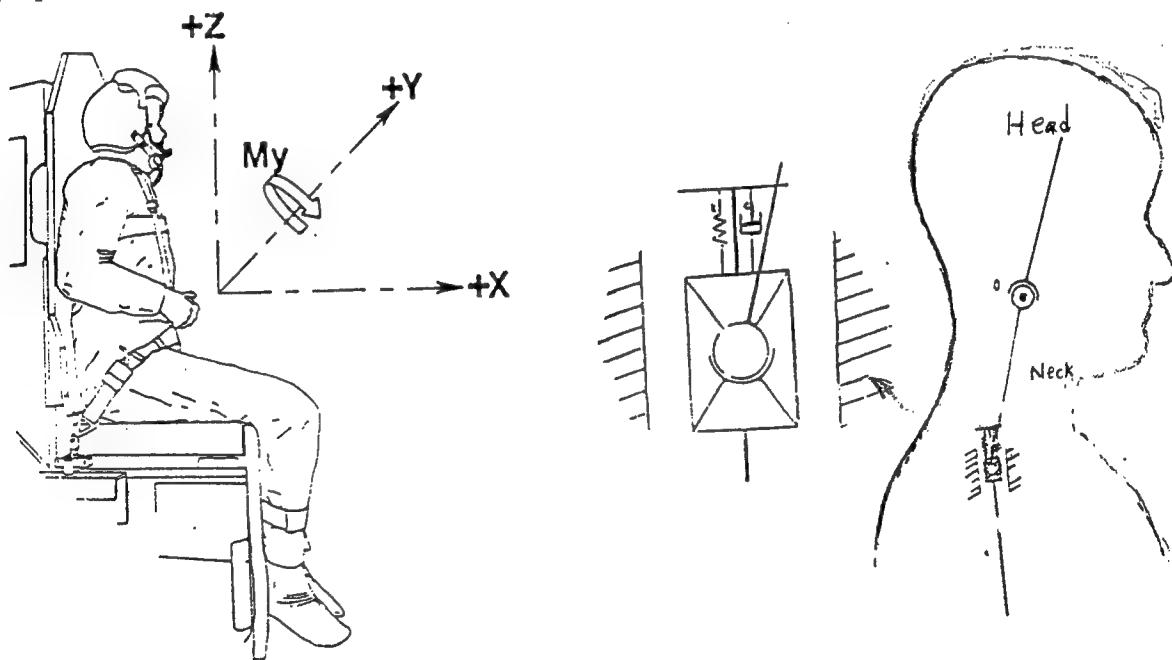
cells were used, the acceleration characteristics of them are shown in Table 1.

Although a total of 15 volunteers ranging from 120 lbs to 218 lb were used, only one subject A-5 (Male- 168lbs) was selected for the analysis reported in this study. This is due to the lack of time and large volume of the data to be processed.

Table 1: Test Cell Characteristics		
Test Cell	Peak Acceleration	Duration
A	4 G	150ms
B	5 G	150ms
C	6 G	150ms
D	7 G	150ms
E	6 G	31ms
F	6 G	64ms
G	6 G	82ms
H	6 G	210ms
I	6 G	250ms

THE MODELING

In this study the Articulated Total Body (ATB) program has been employed. ATB is a rigid body dynamics computer model of the human body. The model is used to predict the kinetic and dynamic response of the human body in different dynamic environments such as a car crash, aircraft pilot ejection, sled test, etc. ATB is the result of a series of modifications made to the original Crash Victim Simulation (CVS) program. In this study, DYNAMAN software, that is a commercialized version of ATB, has also been used for the model preparation or preprocessor of ATB.



In this study a head/neck model consist of three segments namely, Head, Neck, and Upper Torso were used. This is due to the fact that the lower extremity of the subjects were completely restrained. Based on the weight and sex of the subject the physical and geometric information of each segment were taken from GEBOB software. GEBOB program is an

interactive computer program that produces the human and dummy body description data used by ATB model. In this model the head segment is joined to the neck by "Head Pin" (HP) joint and the neck is connected to the upper torso by the "Neck Pin" (NP) joint, see figure 2b. Both joints were of the ball-and-socket type. However, at the neck pin, NP, a viscoelastic slip joint composed of parallel spring and dashpot (Voigt Model) was added. This joint allows relative axial motion between the neck and the upper torso which simulates "compression" at this joint. The spring was represented by linear spring coefficient, K_1 (lb/in) and a quadratic spring coefficient, K_2 (lb/in²), while the dashpot was has only a linear damping coefficient, c (lb-sec/in). Columb friction was assumed zero.

VALIDATION OF THE TEST

The relatively rigid harnessing of the torso and the lower extremity of the subject to the sled leads to the assumption that the acceleration data collected at the chest should be used as the input to the ATB head/neck model. Therefore, the data collected at the mouth piece of the model is the output. One of the objectives of this study is to determine model parameters so that ATB model predicts and duplicate the experimental data collected at the mouth piece. To validate the experiment the test results of human subject B-13 (male 176 lb, test # 4267) of cell C, 6G Amplitude and 150 ms duration, was used. The linear and quadratic coefficients of the spring and the dashpot of the ATB model were varied and the head (at the mouth) acceleration output were compared to the mouth pack acceleration of the test results. After many trial runs and within the time frame of this study, the coefficients were selected and are shown in Table 2. Figures 3, 4 and 5 show the comparison of the test results and the ATB results for the head (mouth pack) accelerations in X, Z and Y directions, respectively. Attention was given to Figure 5, the Y acceleration, where the pattern (peak and valley) of the test results reasonably match with the ATB results. It was observed that the loads and torques at the joints did not changed significantly with variation of these parameters. And, since the main goal of the project was to assess the loads and torques at the joints, the parameters shown in Table 2 were selected.

RESUTLS

The test results of human subject A-5 (male, 168 lbs) in cells A, B, C, D, F, G, H, and

I were analyzed. The test numbers of the cells are 4128, 4147, 4165, 4185, 4226, 4310, 4281, 4298, respectively. The amplitude and time duration of G-y acceleration of the sled in each cell is given in table 1. The chest acceleration of the subject is used as the input to the ATB model. The forces and the torques at the head-pin, HP, and neck-pin, NP, joints in X, Y and Z directions, Fx, Fy and Fz, and Tx, Ty and Tz, respectively, were determined. Figures 6 to 13 show the results of the forces and torques in NP and HP joints. Each figure has two plots of loads and torques for the HP joint and two plots of loads and torques for the NP joint.

As expected, the force in the y direction, Fy is the dominating load at both the neck pin NP and the head pin HP joints. The force Fx in the neck pin NP is almost zero because of the restraining of the chest. However, at the head joint HP, the magnitude of the forces in y and z directions are relatively high compared to that of x. This indicates that the head is moving in the three directions. The torque in the x direction Tx is the dominating load at both the neck-pin, NP, and the head-pin, HP joints. It is important to note that the magnitude of the Tx in NP joint is almost twice as that of the HP joint.

The maximum loads and torques applied to the joints play an important role in neck injuries. Figures 14 depicts the variation of the maximum compressive forces at the NP and the HP joints with respect to the change in amplitude of the acceleration of the test cells. Figure 15 depicts the changes of the maximum torques at the NP and the HP joints versus the change in duration of the acceleration of the test cells. These curves are drawn using a least square curve fitting program.

CONCLUSION

Figure 14 indicates that at the NP joint the maximum force and the maximum torque increases with the increase in the amplitude of the input acceleration. However, the maximum compressive force at HP joint has a minimum. This is due to the fact that the mass inertia of the head has some time lag and back lash with respect to the input acceleration. On the other hand, Figure 15 indicates that the forces and the torques on both joints will increase when the time duration of the input acceleration increases. However, slopes of the curves for the maximum forces and torques is reduced after about 190 ms. This could be due to the fact that muscles responded to the input acceleration and began to absorb some of the load.

Further studies are needed to compare these maximum loads and torques to the human tolerance levels. In addition, for more detailed analysis similar plots should be generated for other subjects. These analyses were beyond the scop of this study and are referred to further investigations.

REFERENCES

1. Guill, F.C. and Herd, G.R., "Aircrew Neck Injuries: A new, or and existing misunderstood phenomenon?", Neck Injury in Advanced Military Aircraft Environment, AGARD, conference Proceedings No. 471, pp,9.1-12, 1990.
2. Knox, F.S., Buhrman, J.R., Perry, C.E., Kaleps, I."Interim head/neck criterion", Armstrong Laboratory Consultation report, WPAFB, OH, Dec., 1991.
3. Miller H.L. and Schneck D.J., "Biodynamic responses to emergency ejections under conditions of added head mass", WPAFB, research contract, 1993.
4. Perry, C.E., "Vertical impact testing of two helmet-mounted night vision systems", AL/CF-SR, 1994-0013

Table 2. Spring and Dashpot Coefficients for the Joints.

1. JOINT TORQUE PROPERTIES:

Table 2a FLEXURAL SPRING PROPERTIES

Joint Name	Linear Coeff.	Quadratic Coeff.	Cubic Coeff.	Energy Coeff.	Joint Stop
NP	450.0	0.000	0.000	0.7000	25.00
HP	450.0	0.000	0.000	0.7000	25.00

Table 2b TORSIONAL SPRING CHARACTERISTICS

Joint Name	Linear Coeff.	Quadratic Coeff.	Cubic Coeff.	Energy Coeff.	Joint Stop
NP	0.000	10.00	0.000	0.7000	35.00
HP	0.000	10.00	0.000	0.7000	35.00

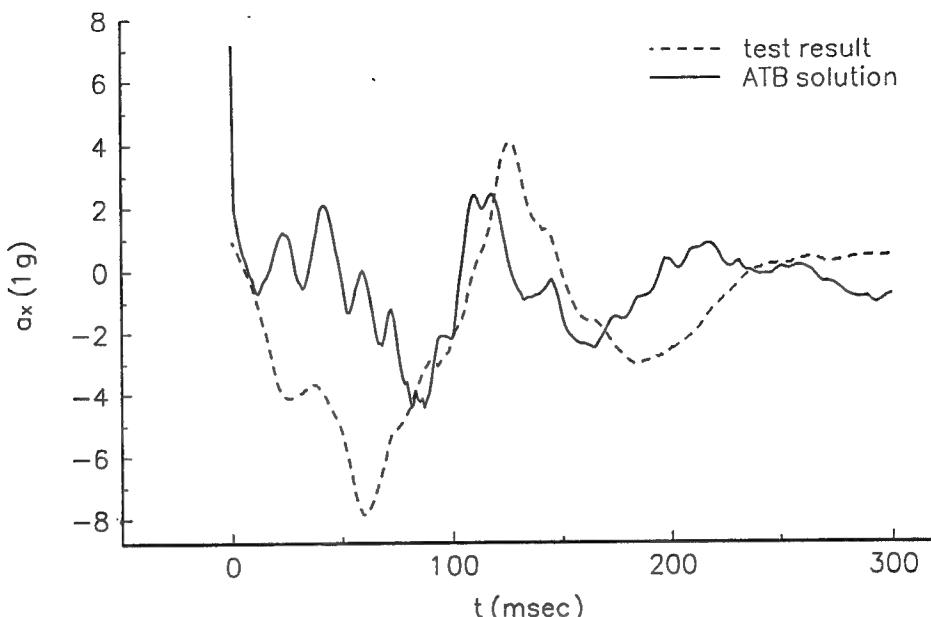
2. JOINT VISCOS PROPERTIES:

Joint Name	Type	Viscous	Coul Fri	Rel Avel	Max Trq	Min Trq	Min Avel	Coef Rest
NP	General	1.000	0.000	30.00	0.000	0.000	0.000	0.000
HP	General	14.9	0.000	30.00	0.000	0.000	0.000	0.000

COMPARISON OF α_x BETWEEN ATB SOLUTION AND TEST RESULT

Data Sheets: 4267hgy2.t21

4267hac1.mot



COMPARISON OF α_z BETWEEN ATB SOLUTION AND TEST RESULT

Data Sheets: 4267hgy2.t21

4267hac1.mot

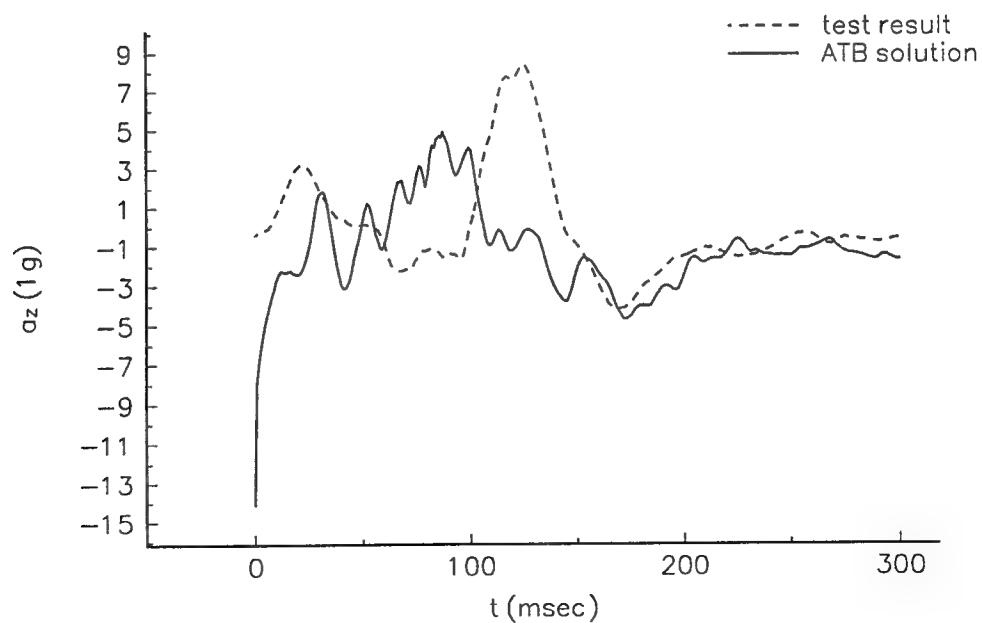


Figure 3 - Validation of α_x

Figure 4 - Validation of α_z

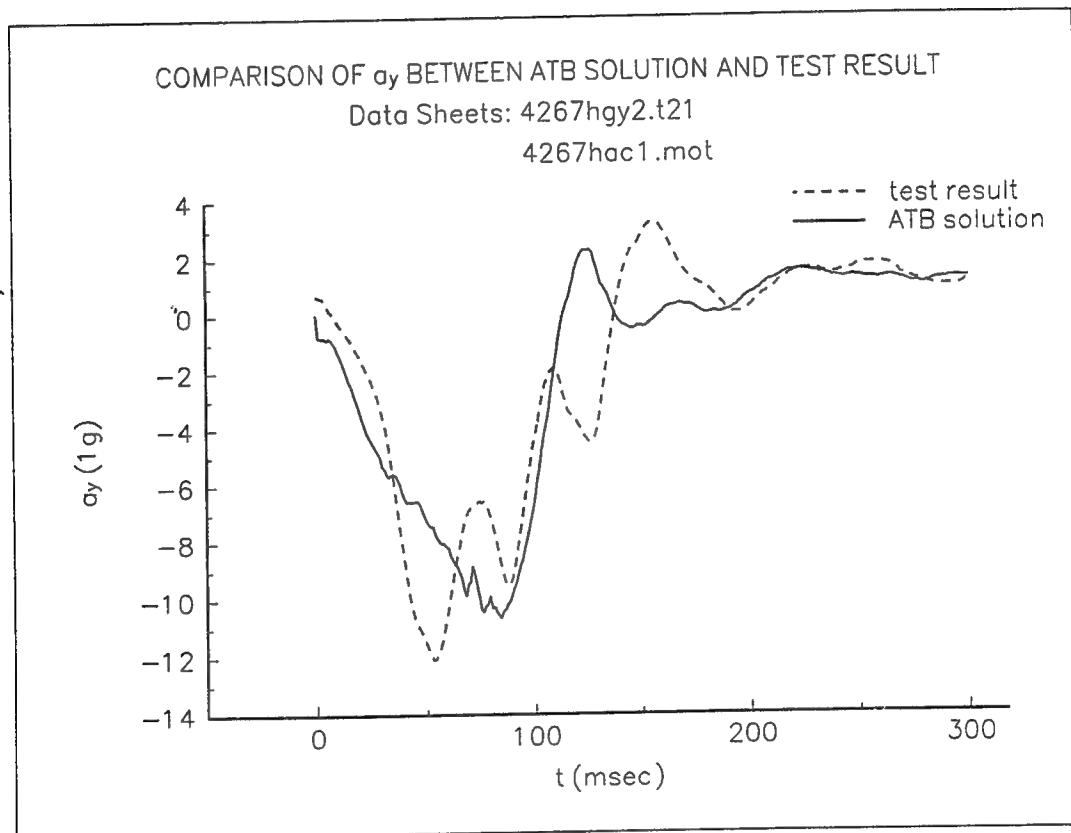


Figure 5 - Validation of a_y

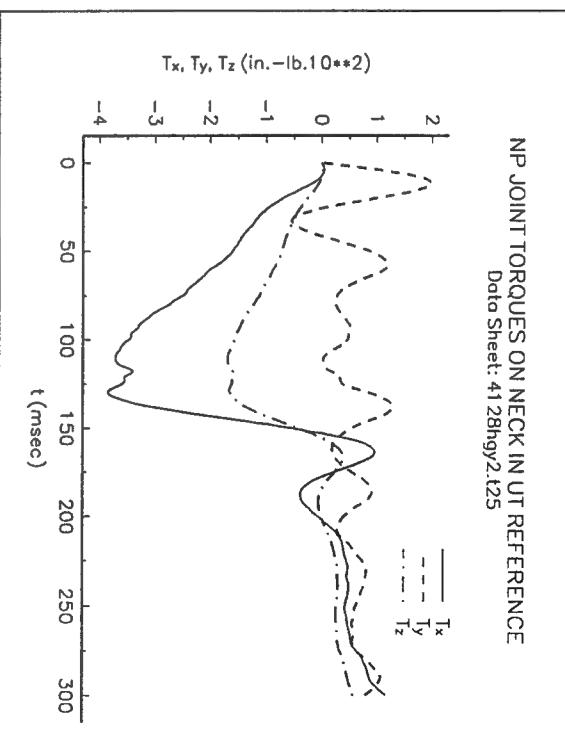
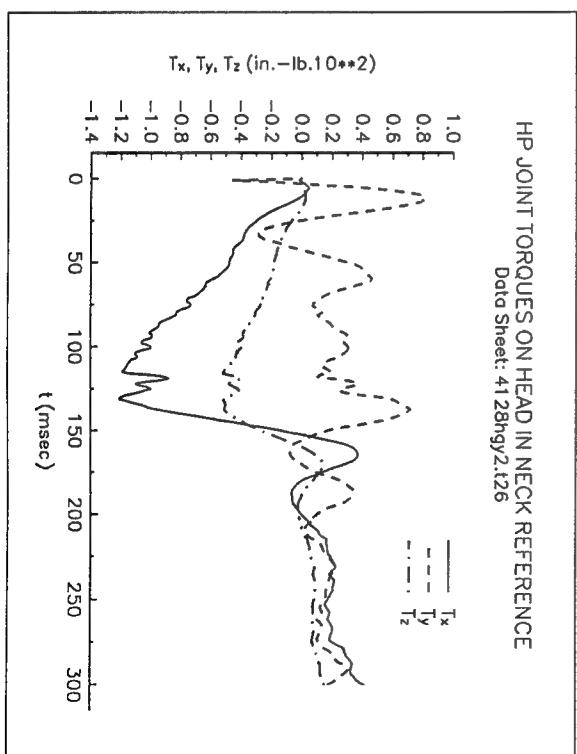
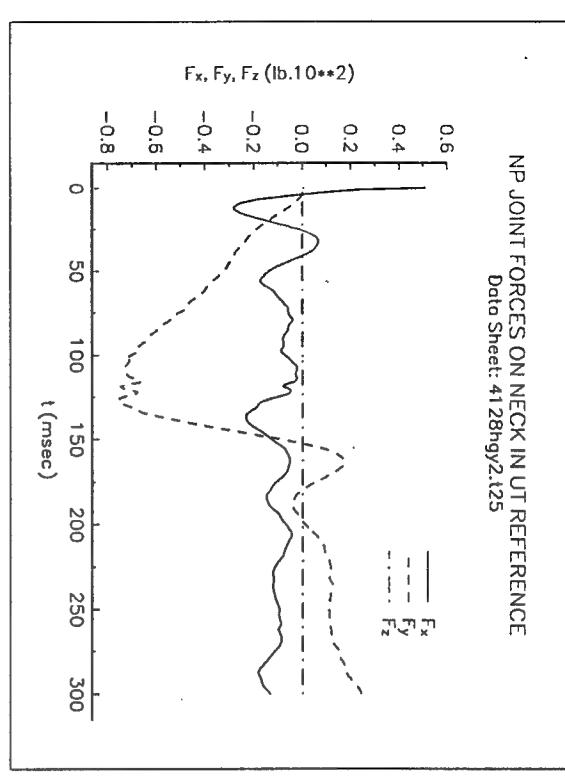
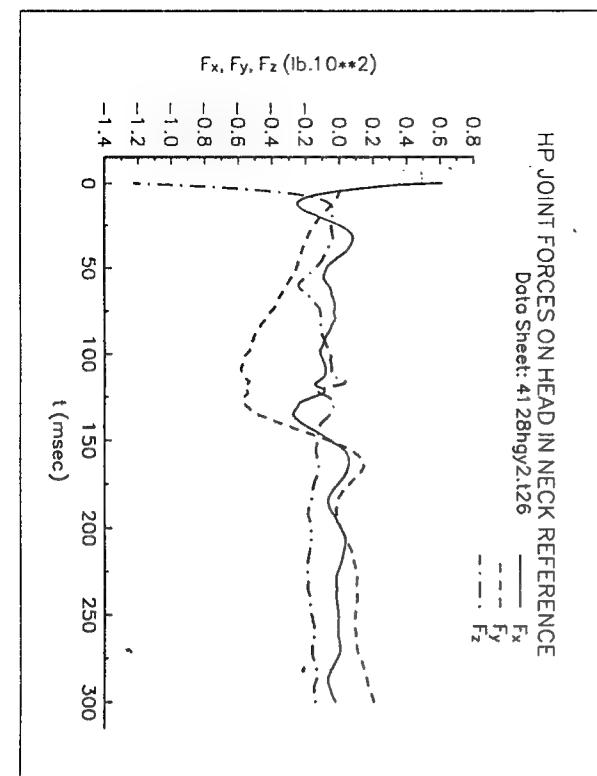


Figure 6 – Forces and Torques at NP and HP Test #4128

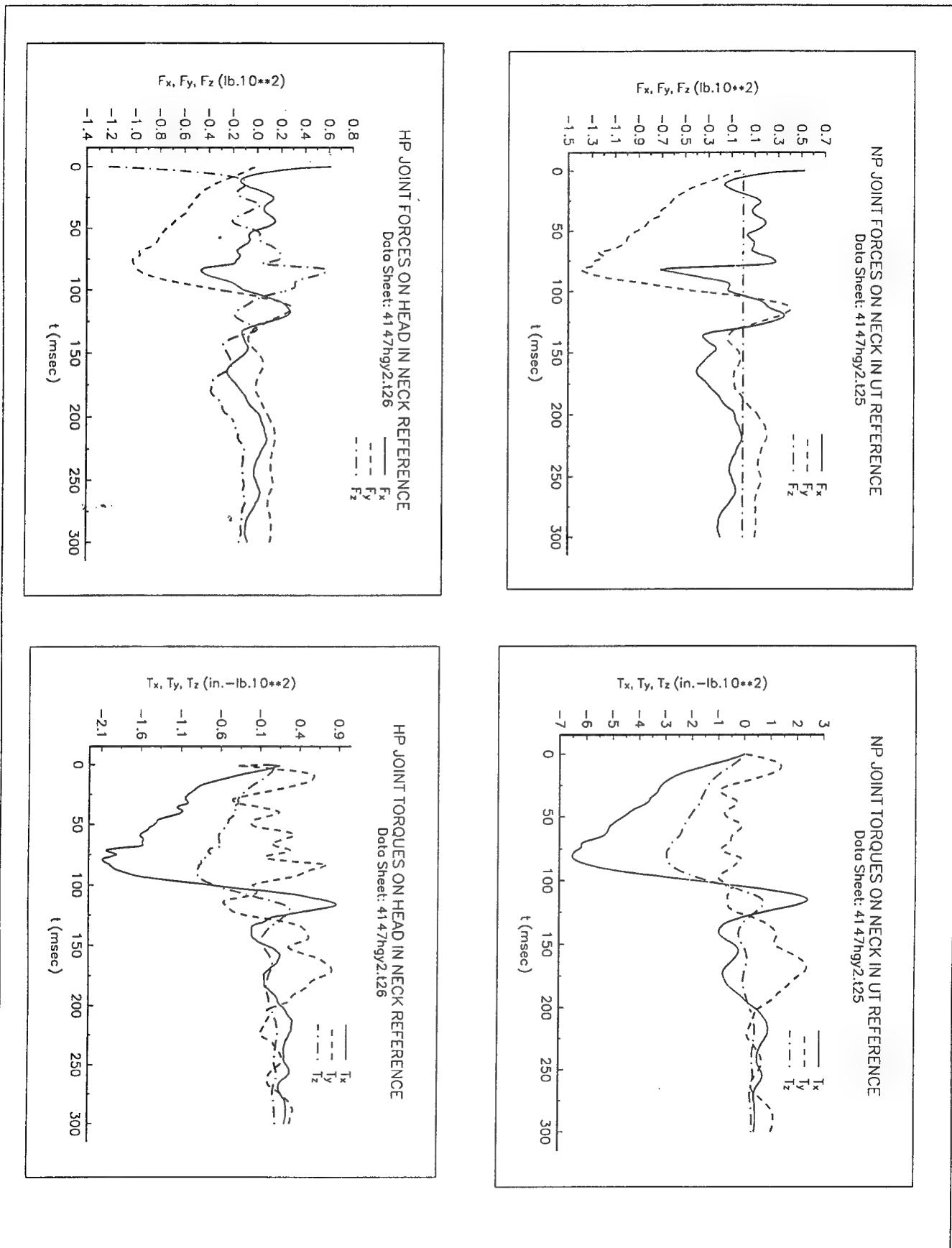


Figure 7 – Forces and Torques at NP and HP Test #4147

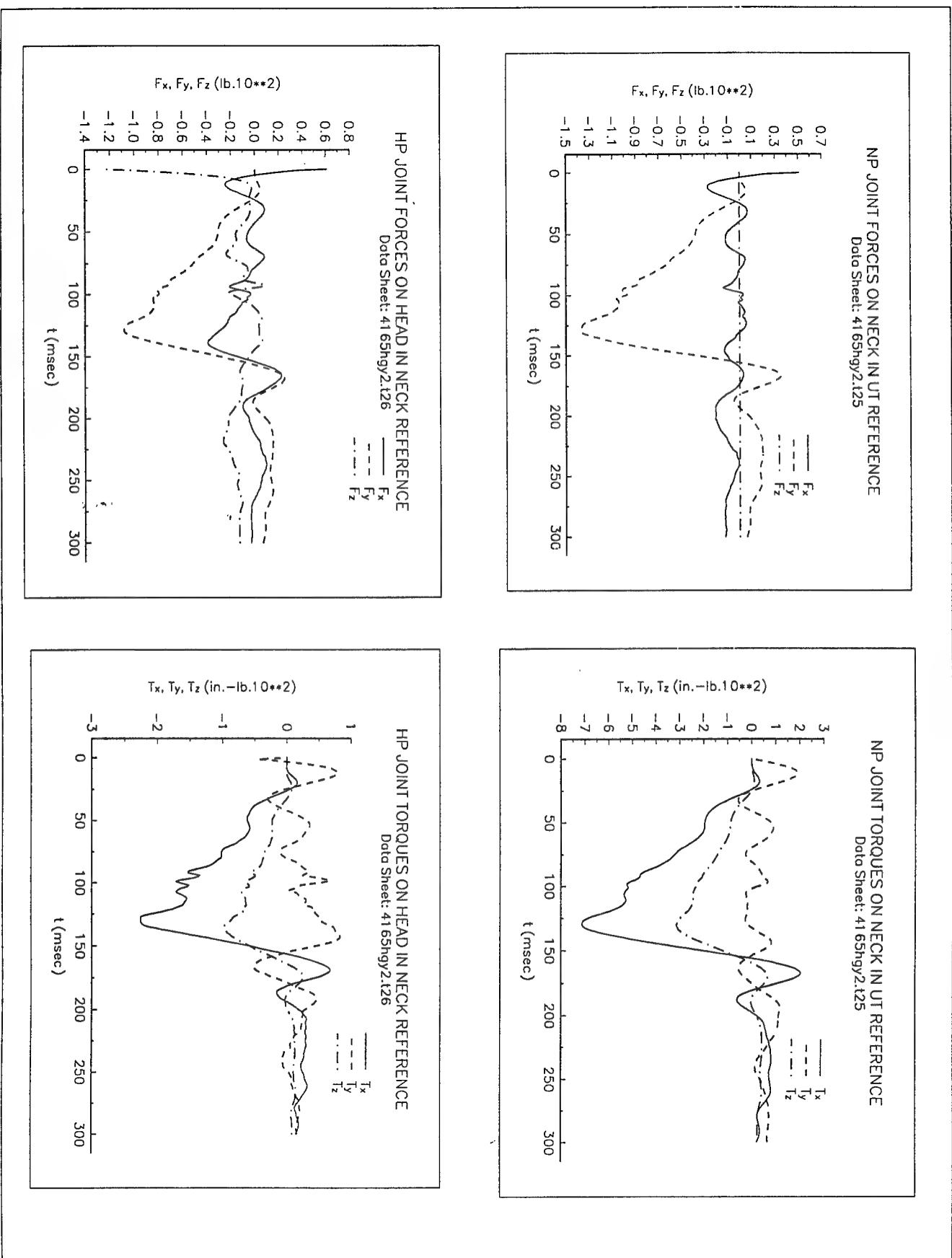


Figure 8 – Forces and Torques at NP and HP Test #4165

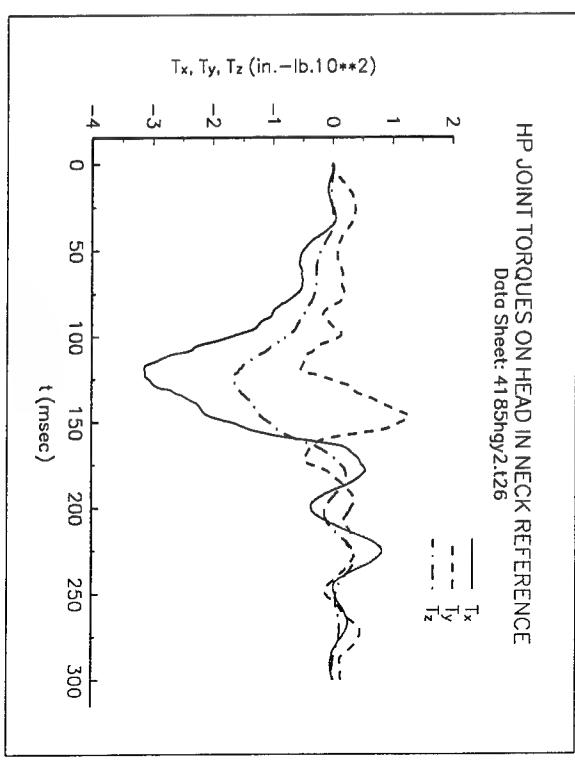
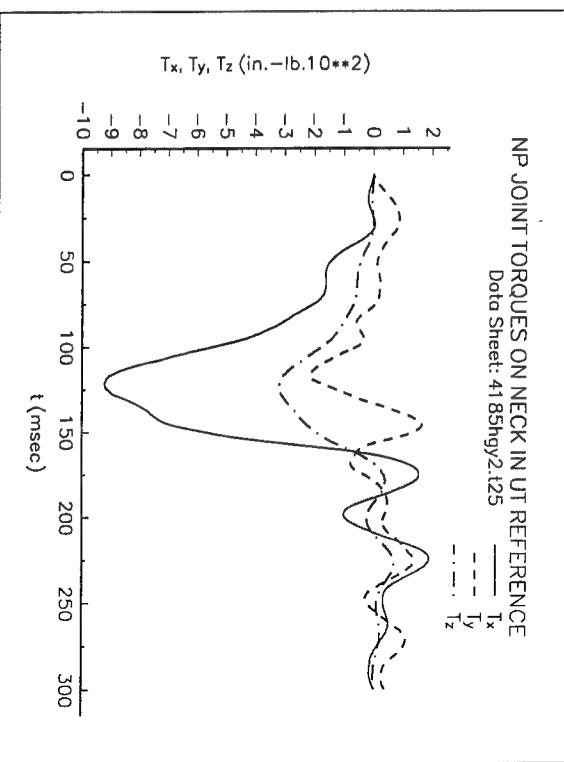
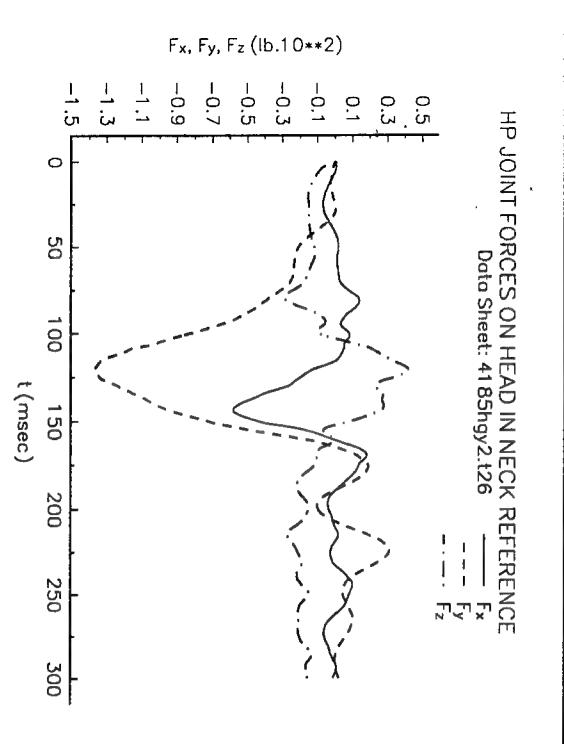
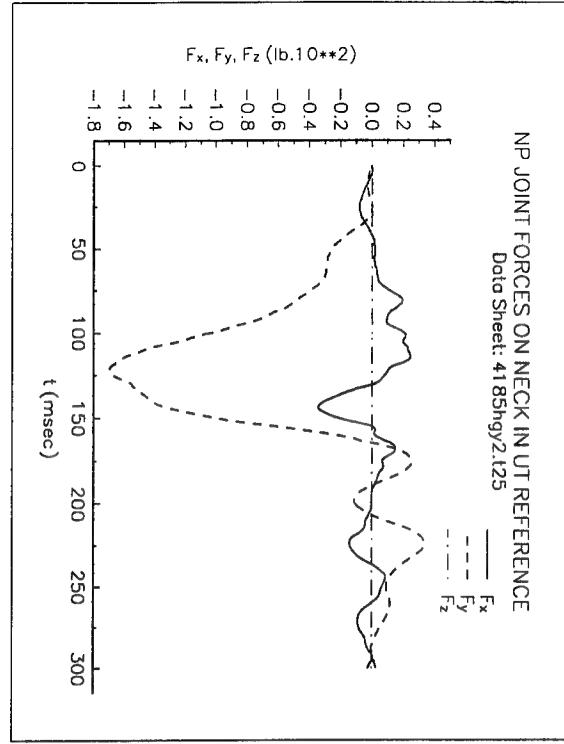


Figure 9 – Forces and torques at NP and HP Test #4185

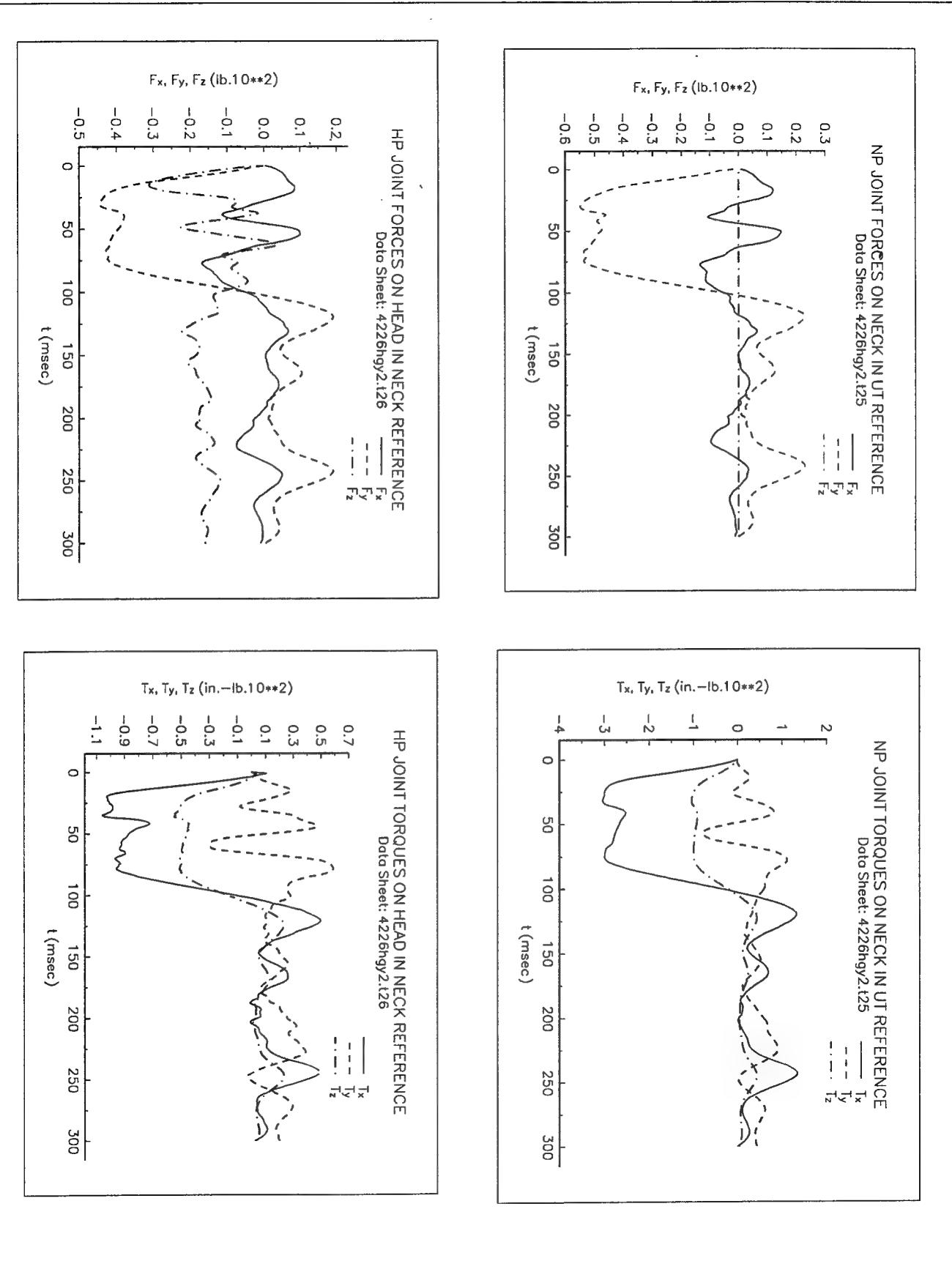


Figure 10 – Forces and torques at NP and HP Test #4226

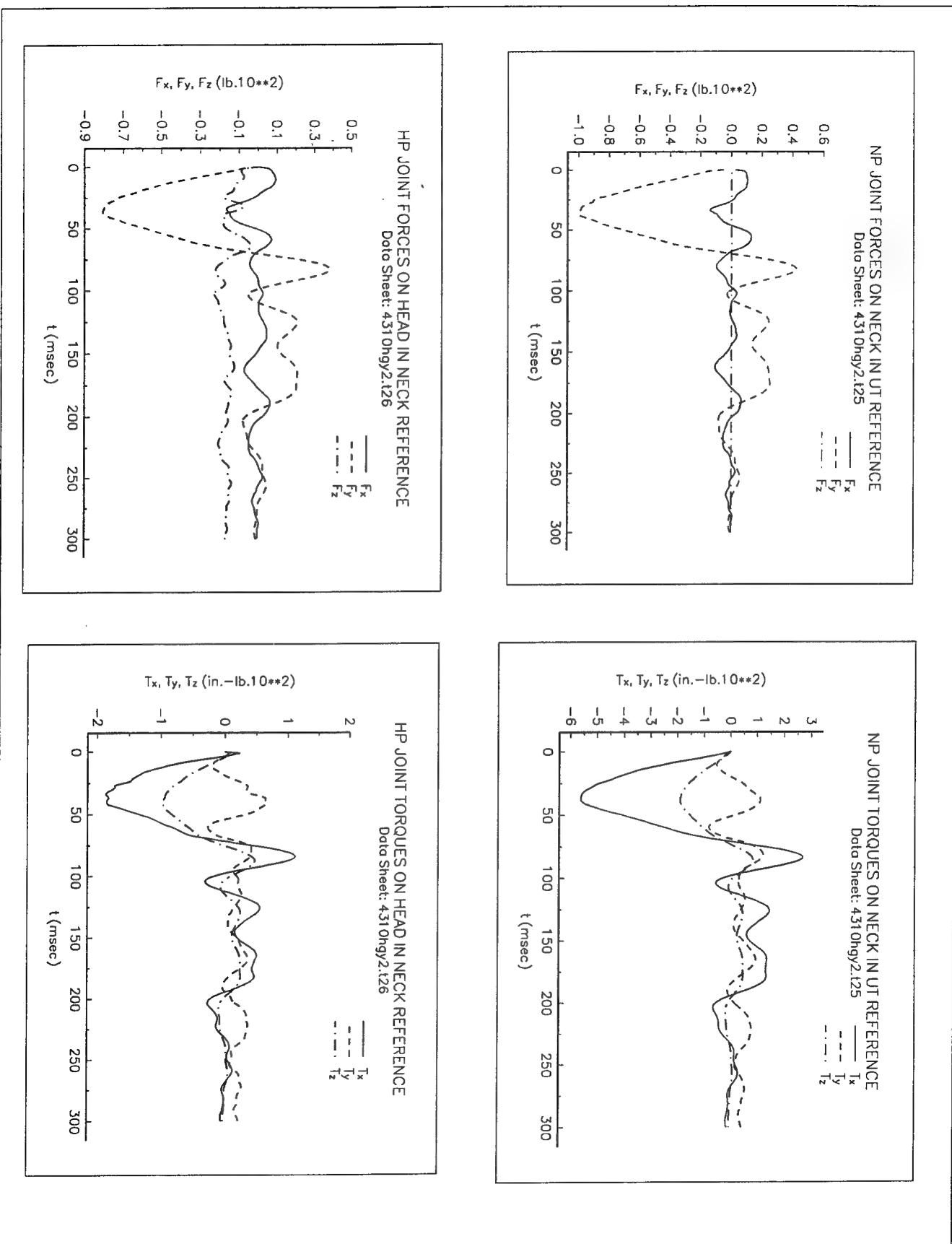


Figure 11 – Forces and Torques at NP and HP Test #4310

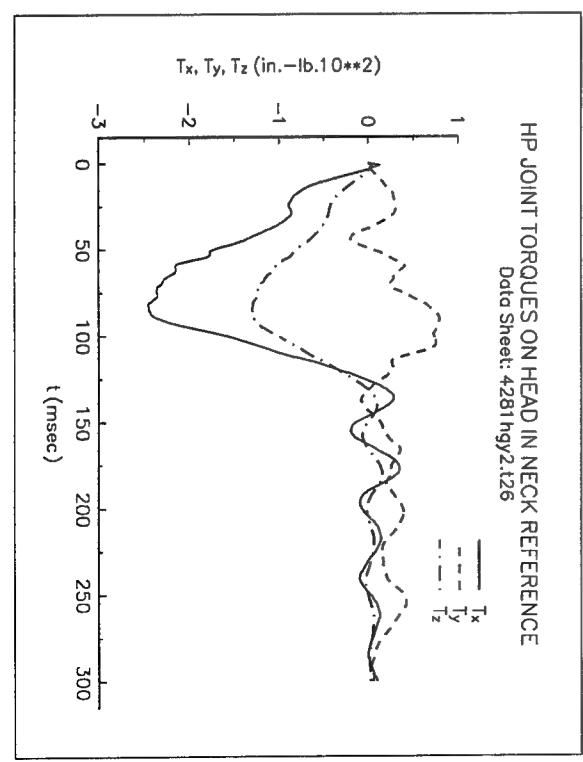
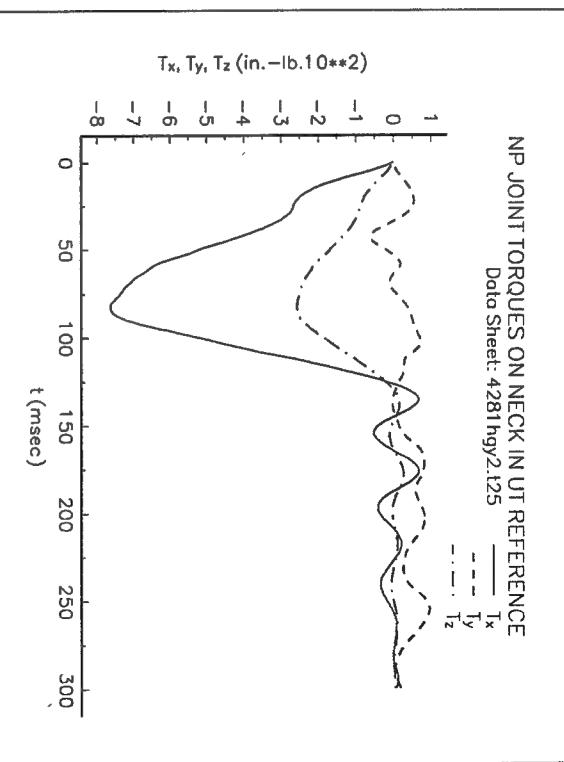
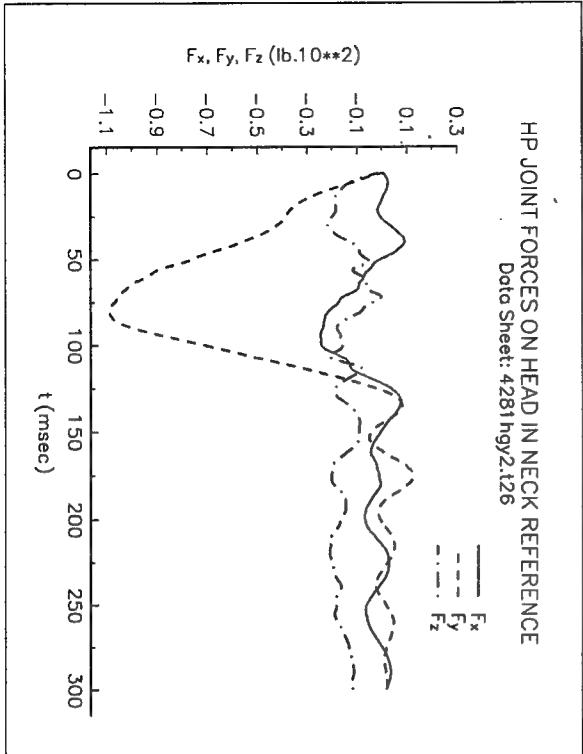
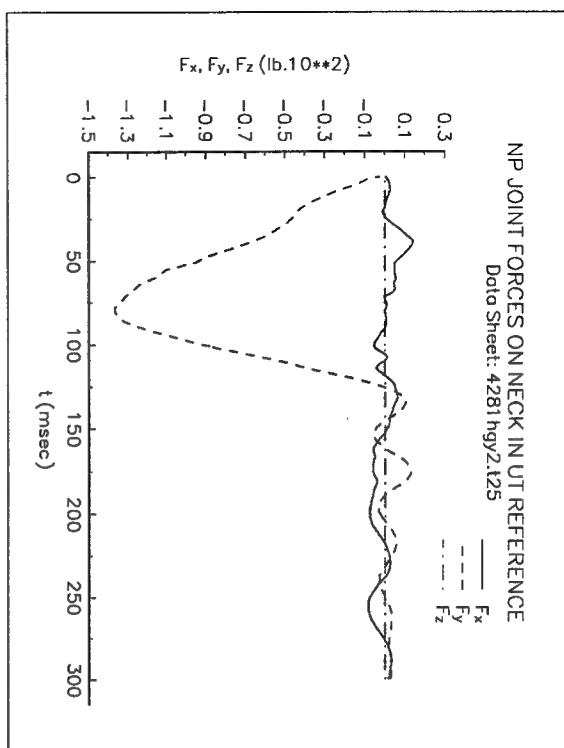


Figure 12 – Forces and Torques at NP and HP Test #4281

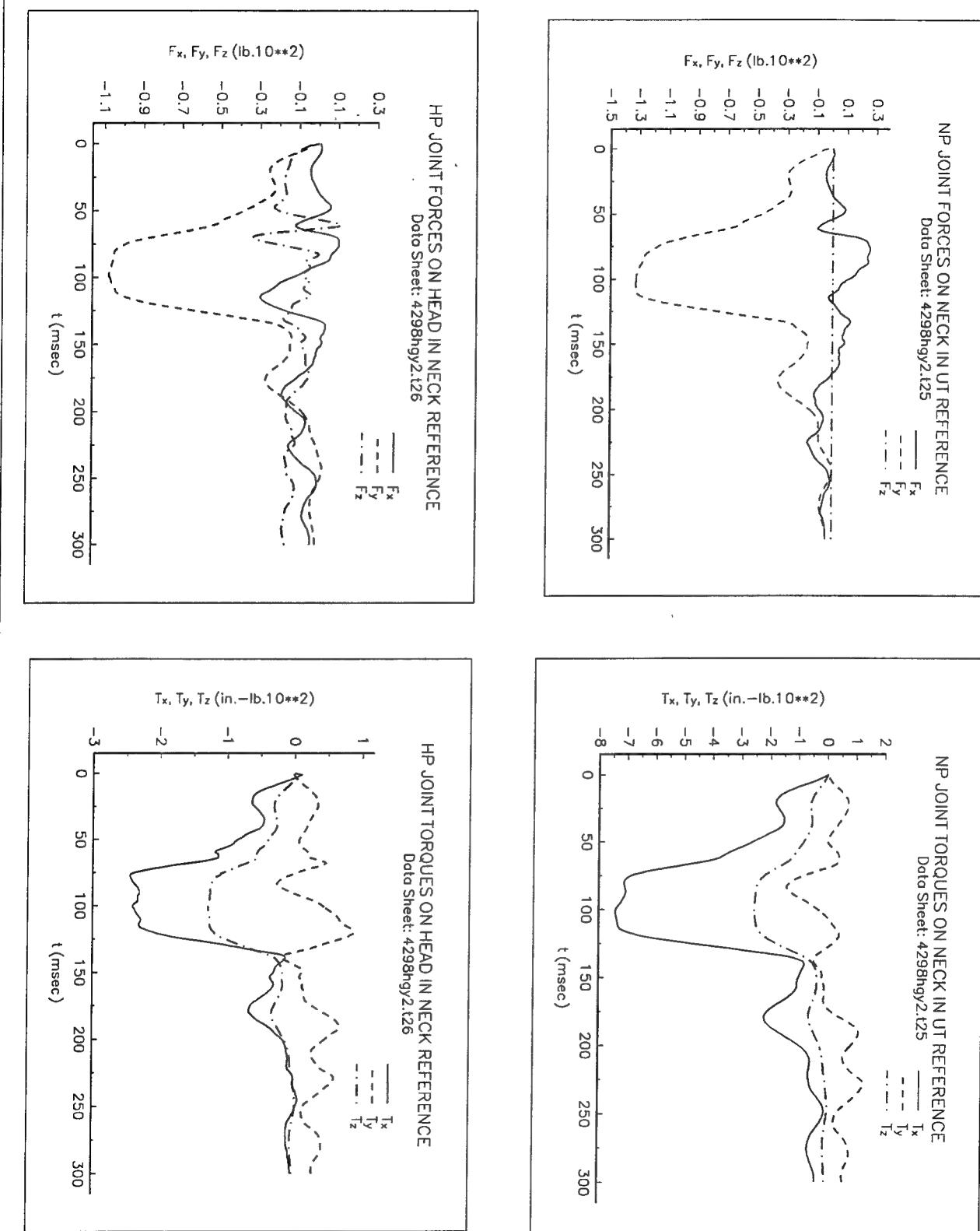
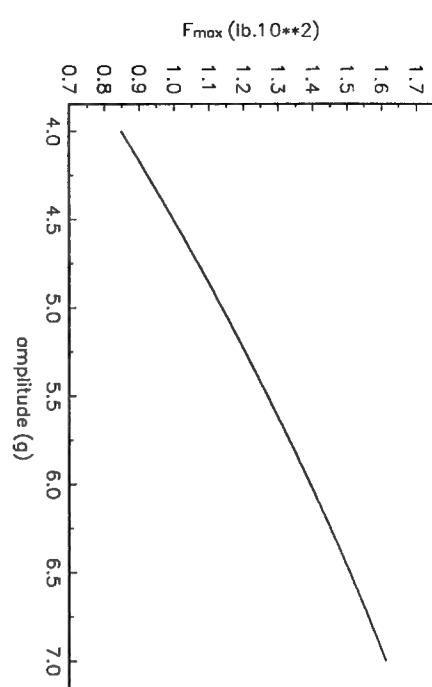


Figure 13 – Forces and Torques at NP and HP Test #4298

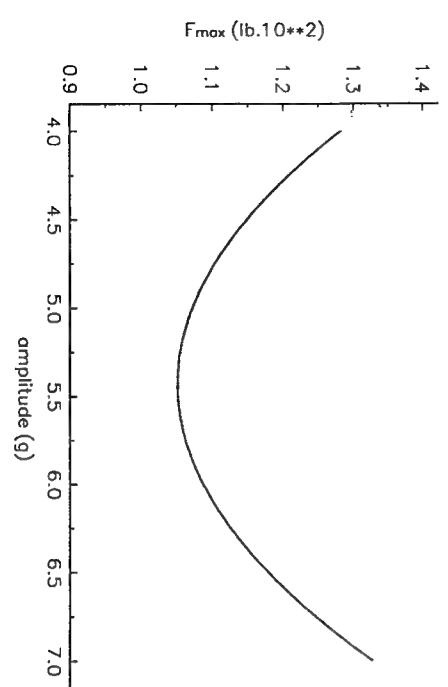
MAXIMUM COMPRESSIVE NP JOINT FORCE—AMPLITUDE DIAGRAM

Data Sheet: fnp.max



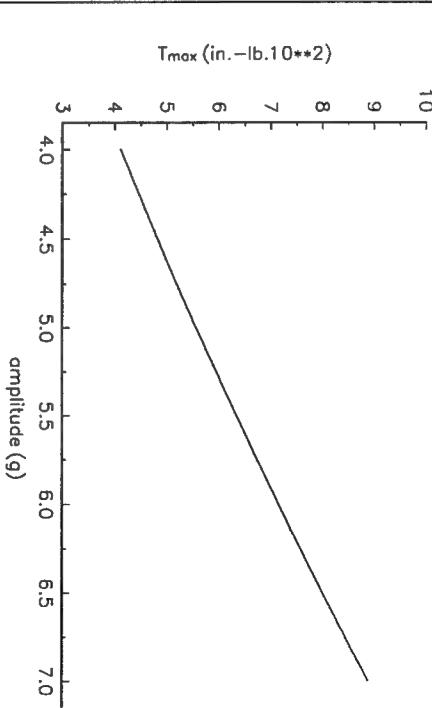
MAXIMUM COMPRESSIVE HP JOINT FORCE—AMPLITUDE DIAGRAM

Data Sheet: fohp.max



MAXIMUM NP JOINT TORQUE—AMPLITUDE DIAGRAM

Data Sheet: tnp.max



MAXIMUM HP JOINT TORQUE—AMPLITUDE DIAGRAM

Data Sheet: tohp.max

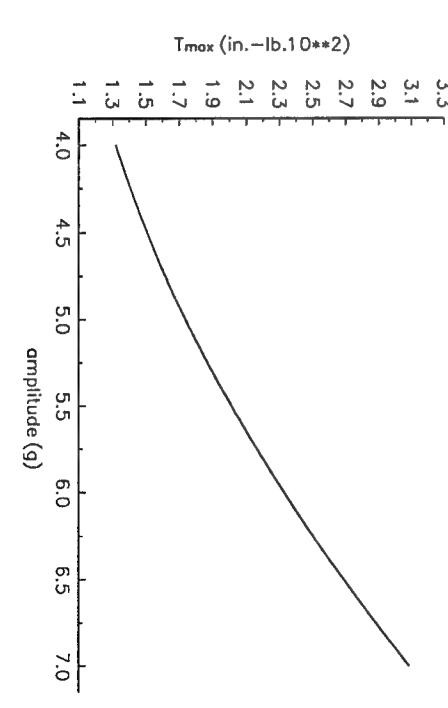


Figure 14 – Maximum Force and Torque vs. Amplitude

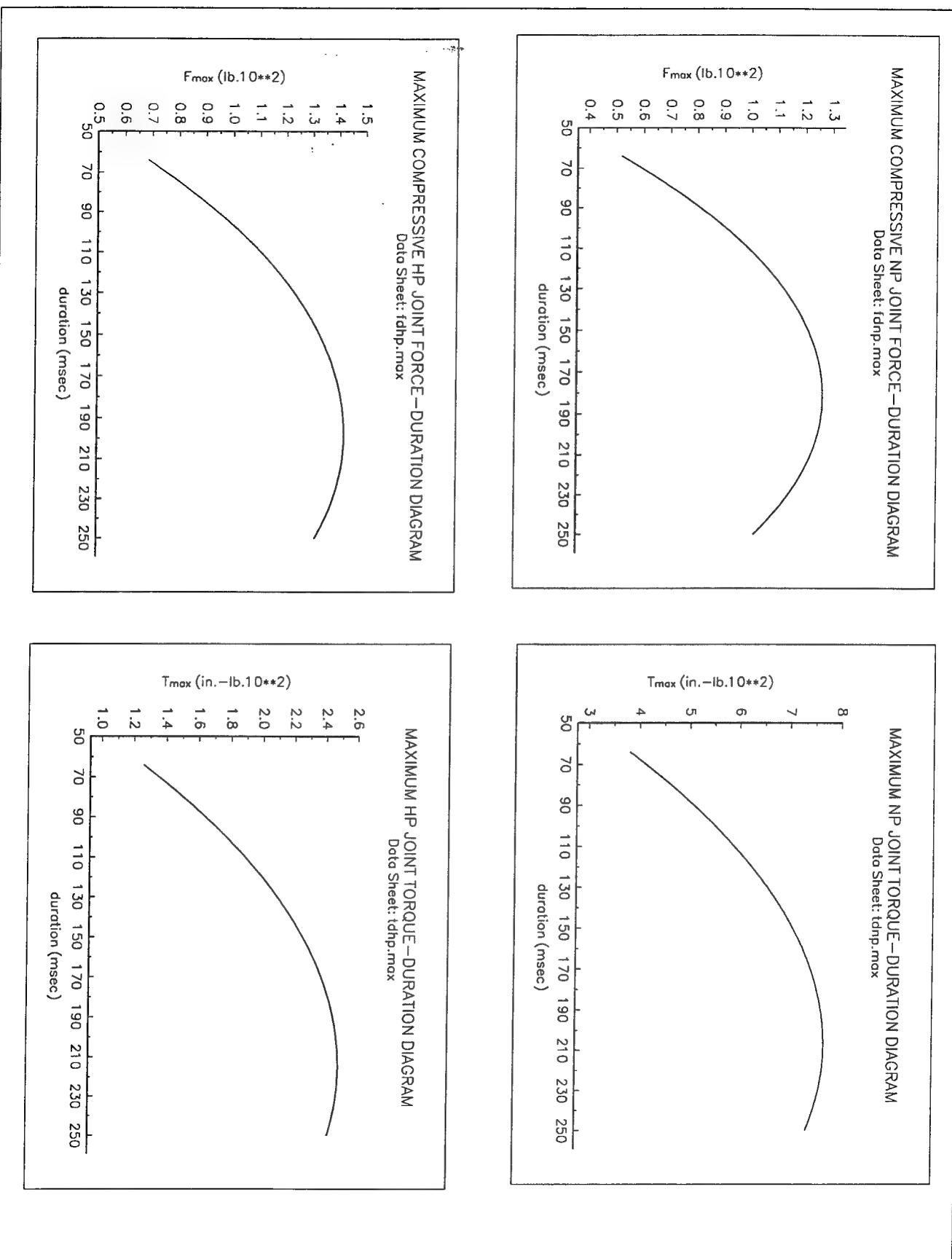


Figure 15 – Maximum Force and Torque vs. Duration

HALF-LIFE STUDIES IN RANCH HAND VETERANS

Kandasamy Selvavel
Associate Professor
Department of Mathematics
and
Computer Science

Claflin College
700 College Avenue
Orangeburg, SC 29115

Final Report for:
Summer Faculty Research Program
Armstrong Laboratory

Sponsored by:

Air force Office of Scientific Research
Bolling Air Force Base, DC

and

Armstrong Laboratory
Brooks Air Force Base
San Antonio, TX

September 1995

HALF-LIFE STUDIES IN RANCH HAND VETERANS

Kandasamy Selvavel

Associate Professor

Department of Mathematics and Computer Science

Claflin College

Abstract

Half-life of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) in veterans of Operation Ranch Hand was studied using measurements from serum collected in 1982, 1987 and 1992. In our study we estimated the half-life using two measurements per subject and three measurements per subject. These half-lives were corrected by finding the regression effect towards the mean. Then the bias, variance and mean square error (MSE) of the estimators were computed and compared for various cases. It was noticed that for large sample unbiased estimators were admissible when we took two measurements per subject. In general this result doesn't hold for small sample study. Furthermore, we studied the mean square error as a function of truncation point. We fixed the second truncation point and moved the first one towards the second one. It is observed that the difference in mean square errors between unbiased and biased estimators changes sign. That is, when the truncation points are closed to each other unbiased estimator has a smaller mean square error than that of biased estimator. On the other hand when the first truncation point moved away from the second point, mean square error of the biased estimator was getting bigger than that of the unbiased estimator.

1. INTRODUCTION

The investigation tetrachlorodibenzo-p-dioxin (TCDD) half-life of Vietnam veterans of Operation Ranch Hand was first studied by Pirkle et al. (1989). In their study, they selected thirty six pairs of observations with at least 10 parts per trillion (10 ppt) concentration of dioxin, and used the standard half-life equation derived from first-order kinetics to estimate the half-life. Using a background correction of 4 ppt, they obtained median half-life as 7.1 years with a 95% confidence interval of 5.8-9.6 years. Michalek et al. (1992) studied the relationship between TCDD half-life and percentage of body fat and changes in percentage of body fat of the same group of 36 Air Force veterans that were considered by Pirkle et al. (1989). The repeated measures analysis of a covariance model

$$y_{ij} = \beta_0 + \tau_i + \beta_1 t_{ij} + \epsilon_{ij},$$

was used in their study, where y_{ij} is the natural logarithm of the j th TCDD background corrected measurement of the i th individual, t_{ij} is the number of years between the i th and first TCDD measurements, β_1 is the common decay rate, $(\beta_0 + \tau_i)$ is the intercept for the i th individual, and ϵ_{ij} is the residual error term for measurement y_{ij} . The half-life corresponding to this model was estimated as 7.68 years with a 95% confidence interval of 6.2-10.1 years. Moreover, it was found that the half-life has only a marginal relationship with percentage body fat and that there was no significant relationship between half-life and relative change in percentage of body fat. Recently, Wolfe et al. (1994) estimated the half-life as 11.6 years using 337 Ranch Hand veterans.

We also assume the first order kinetics model

$$C_t = C_0 e^{-\lambda t} \quad (1.1)$$

holds, where C_t is the concentration t years after exposure, C_0 is the initial concentration, and λ is the unknown decay rate. The half-life corresponding to this model is given by $H_L = \ln(2)/\lambda$. Equation (1.1) can be written as $\log C_t = \log C_0 - \lambda t$, which motivates us to use a repeated measures linear model

$$y_{ij} = \mu + \tau_i + \beta t_{ij} + \varepsilon_{ij}, \quad (1.2)$$

for $i=1,2,\dots,n$ and $j=1,2,\dots,k$, where i indexes n subjects and j indexes k measurements per subject, y_{ij} represents the logarithm of the j th measurement on the i th subject, $-\beta$ is the common decay rate (λ), $t_{ij} = t_i + (j-1)\Delta$ is the time between exposure and the j th measurement, τ_i is the fixed effect for the i th subject and ε_{ij} is the residual error term for y_{ij} . In our study we estimate the half-life by finding the regression effect on the decay rate. In the first approach, we first estimate the regression effect using two time points, and it is found as the products k , σ and $(1-p)$, where $k = \phi(c)/(1-\phi(c))$. Here, c is the standardized cut point and ϕ , ϕ are the standard normal density and cumulative distribution function respectively. In the second approach we estimate the regression effect using three time points. Then the corresponding half-lives are calculated by adjusting these regression effect.

Next we study the bias, variance and mean square error (MSE) of the estimators. Under AR(1) model expressions for variance of $\hat{\beta}$ are given for two measurements per subject and three measurements per subject. More over, mean square errors of $\hat{\beta}$ are computed and compared for unbiased and biased estimators for various cases. This is summarized and given in table 3.1. Further more, we study the mean square error as a function of truncation point.

2. REGRESSION EFFECT ON DECAY RATE

In this section we consider a simple repeated measures linear model

$$y_{ij} = \mu + \tau_i + \beta t_{ij} + \varepsilon_{ij}, \quad 1 \leq i \leq n, \quad 1 \leq j \leq 3,$$

and estimates the decay rate ($-\beta$) by considering the regression effect. Here, y_{ij} is the natural logarithm of the j th dioxin measurement of the i th individual, t_{ij} is the number of years between the j th and first dioxin measurements, β is the slope of the regression line and ε_{ij} is assumed to be normal with mean 0 and standard deviation σ for all i and j . Now we assume that $Y = (Y_1, Y_2, Y_3)'$ is normally distributed with mean vector $\mu = (\mu + \tau, \mu, \mu - \tau)'$ and covariance matrix

$$\Sigma = \sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & \rho \\ \rho^2 & \rho & 1 \end{bmatrix}$$

set C be the cut point of the second measurement

Then it can be easily shown that

$$E(\hat{\beta} | y_2 > c) = \frac{-\tau}{\Delta} \left[1 + \frac{3(n-1)s_t^2}{2n\Delta^2} \right],$$

where $\Delta = t_{i2} - t_{i1} = t_{i3} - t_{i2}$ and $s_t^2 = \sum \sum t_{ij}^2$.

If all the subjects were exposed at the same time we would have an unbiased estimator of the decay rate. In this special case there is no regression effect.

Using 240 Ranch Hand veterans with dioxin measurements taken in 1982, 1987 and 1992, we have $n=240$, $\Delta=5$, $c=10$ and $s_t^2=2.55$. Hence, the proportion of the regression effect on the decay rate is 0.86956. Then using the uncorrected decay rate 0.776, the corrected decay rate is given by 0.08924, and the corresponding half-life is 7.77 years.

Next we consider the situation with two time points. Let X_1 and X_2 be the random variables corresponding to the logarithms of 1987 and 1992 dioxin concentration of the Ranch Hand veterans respectively. We assume that $X=(X_1, X_2)$ is normally distributed with mean vector $\mu=(\mu, \mu-\tau)$, $\tau > 0$, and covariance matrix

$$\Sigma = \sigma^2 \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$$

Let ΔT be the time from 1987 dioxin measurement until 1992 dioxin measurement in unit of years, and c_0 be the left truncation point of logarithm of 1987 dioxin concentration. Then the regression effect on estimating decay rate of dioxin concentration by truncating the first time point is given by

$$\{k_0\sigma(1-\rho)\}/\Delta T, \quad (2.1)$$

where $k_0 = \phi(C_0)/[1 - \Phi(C_0)]$,

$C_0 = (c_0 - \mu)/\sigma$, and ϕ and Φ are the standard normal density and cumulative distribution function respectively. It can be easily shown that the same result holds when we estimate the regression effect of the decay rate of the dioxin concentration from 1982 to 1987 by truncating the second time point.

Furthermore, if we truncate on both time points the regression effect on estimating decay rate of dioxin concentration is given by

$$\frac{\sigma(1-\rho)}{P(X>C_1, X>C_2)}[\phi(C_1)-\phi(C_2)+\Phi(C_2)\Phi\left(\frac{C_1-\rho C_2}{\sqrt{1-\rho^2}}\right)-\Phi(C_1)\Phi\left(\frac{C_2-\rho C_1}{\sqrt{1-\rho^2}}\right)],$$

where $C_1=(c_1-\mu)/\sigma$ and $C_2=(c_2-\mu+\tau)/\sigma$.

Suppose ΔT , the time from 1987 dioxin concentration until 1992 dioxin measurement. Using the regression approach considered in Michalek (1992), the uncorrected decay rate is given by 0.095031. Therefore, the corresponding uncorrected half-life is $\ln(2)/0.095031=7.3$ years. By taking C_0 as the 0.40 quantile of 1987 dioxin distribution, we have $k_0 = 0.64383$. Using Senn and Brown (1985), we get $\rho=0.89204$. Taking the standard deviation of the logarithm of 1987 dioxin concentration $\sigma=1.054023$, and using (2.1) the regression effect is given by 0.014389. Therefore, the true decay rate is 0.0806397 and hence the corresponding corrected half-life is given by $\ln(2)/0.0806397=8.6$ years.

Using the same approach with 1982 and 1987 dioxin concentration we can easily show that the corrected half-life obtained by truncating on both time points is given by 9.0 years.

3. VARIANCE, BIAS AND MEAN SQUARE ERROR OF THE ESTIMATOR

In this section we give expressions for the bias, variance and mean square error of $\hat{\beta}$ under AR(1) assumption. Using 1982, 1987 and 1992 dioxin measurements we compute the bias, variance and mean square error for various estimates of β . Then using these computations we compare the estimates of β for two measurements per subject and three measurements per subject.

First we look at the case where two measurements per subject are available for the study. Under AR(1) assumption $\hat{\beta}$, $\text{var}(\hat{\beta})$ and bias for 1982 and 1987 dioxin measurements are given by

$$\hat{\beta} = \frac{1}{\sum(\Delta_i)^2} \sum \Delta_i (y_{i2} - y_{i1})$$

$$\text{var}(\hat{\beta}) = \frac{1}{[\sum(\Delta_i)^2]^2} \sum (\Delta_i)^2 \{ \text{var}(y_{i1}) + \text{var}(y_{i2}) - 2\text{cov}(y_{i1}, y_{i2}) \}$$

and

$$\text{bias} = \frac{\tau}{\Delta_i} - \frac{1}{\sum(\Delta_i)^2} \sum \Delta_i \{ E(y_{i2}) - E(\hat{y}_{i1}) \},$$

where $\Delta_i = 5$. For convenience in notation $Y_{i1} > a_{i1}$, $Y_{i2} > a_{i2}$ is left out. Let X_{i1} and X_{i2} be the standardized random variables corresponding to Y_{i1} and Y_{i2} respectively. Then the expected values of Y_i 's, variances of Y_i 's and covariance between Y_{i1} and Y_{i2} can be calculated using the following expressions that are derived from Tallis (1961).

$$E(X_{i1}) = \frac{1}{\alpha_i} [\phi(a_{i1})\Phi(-A_{i,12}) + \rho\phi(a_{i2})\Phi(-A_{i,21})]$$

$$E(X_{i2}) = \frac{1}{\alpha_i} [\rho\phi(a_{i1})\Phi(-A_{i,12}) + \phi(a_{i2})\Phi(-A_{i,21})]$$

$$E(X_{i1}^2) = 1 + \frac{1}{\alpha_i} [a_{i1}\phi(a_{i1})\Phi(-A_{i,12}) + \rho^2 a_{i2}\phi(a_{i2})\Phi(-A_{i,21}) + \rho(1-\rho^2)\phi(a_{i2}, a_{i1}; \rho)]$$

$$E(X_{i2}^2) = 1 + \frac{1}{\alpha_i} [\rho^2 a_{i1}\phi(a_{i1})\Phi(-A_{i,12}) + a_{i2}\phi(a_{i2})\Phi(-A_{i,21}) + \rho(1-\rho^2)\phi(a_{i2}, a_{i1}; \rho)]$$

and

$$E(X_{i1}X_{i2}) = \rho + \frac{1}{\alpha_i} [\rho a_{i1}\phi(a_{i1})\Phi(-A_{i,12}) + \rho a_{i2}\phi(a_{i2})\Phi(-A_{i,21}) + (1-\rho^2)\phi(a_{i2}, a_{i1}; \rho)],$$

where $\alpha = P(X_{i1} > a_{i1}, X_{i2} > a_{i2})$ with (X_{i1}, X_{i2}) , the bivariate standard normal random variables with correlation coefficient ρ , $\phi(x, y; \rho)$, the bivariate standard normal density function with correlation coefficient ρ , and

$$a_{i1} = \frac{c_1 - y_{i1}^*}{\sigma}, \quad a_{i2} = \frac{c_2 - y_{i2}^*}{\sigma},$$

$$A_{i,12} = \frac{a_{i2} - \rho a_{i1}}{\sqrt{1-\rho^2}} \quad \text{and} \quad A_{i,21} = \frac{a_{i1} - \rho a_{i2}}{\sqrt{1-\rho^2}}.$$

Here, $\phi(a)$ is the standard normal density function and $\Phi(a) = P(Z < a)$ with Z being a random variable with standard normal density function.

Note that when the standardized cut points are same $\hat{\beta}$ is an unbiased estimator of β . In this special case bias is zero and variance of $\hat{\beta}$ is given by

$$\text{var}(\hat{\beta}) = \frac{\sigma^2}{25n^2} \sum \{ \text{var}(X_{i1}) + \text{var}(X_{i2}) - 2\text{cov}(X_{i1}, X_{i2}) \},$$

where

$$E(X_{i1}) = E(X_{i2}) = \frac{(1+\sigma)}{\alpha_i} \phi(a_{i1})\Phi(-A_{i,12})$$

$$E(X_{i1}^2) = E(X_{i2}^2) = 1 + \frac{1}{\alpha_i} [(1+\rho^2)a_{i1}\phi(a_{i1})\Phi(-A_{i,12}) + \rho(1-\rho^2)\phi(a_{i1}, a_{i1}, \rho)]$$

and

$$E(X_{i1}X_{i2}) = \rho + \frac{1}{\alpha_i} [2\rho a_{i1}\phi(a_{i1})\Phi(-A_{i,12}) + (1-\rho^2)\phi(a_{i1}, a_{i1}, \rho)].$$

Here $\phi(x,y;\rho)$ is a bivariate standard normal density function with correlation coefficient ρ .

Similar expressions can be derived for 1987, 1992 and 1982, 1992 dioxin measurements.

Table 3.1

<u>bias</u>	<u>variance</u>	<u>MSE</u>
3.6093086E-03	2.3297681E-05 32-8	3.6324790E-05
Year 82-87, n=365, $\rho=0.88497$, $a_{i1} \neq a_{i2}$		
1.705659E-04	2.4414508E-05	2.4443601E-05
Year 87-92, n=240, $\rho=0.82286$, $a_{i1} \neq a_{i2}$		
1.0548597E-02	5.6190886E-05	1.6746379E-04
Year 87-92, n=213, $\rho=0.82286$, $a_{i1} = a_{i2}$		
3.6289920E-04	6.1123183E-05	6.1254879E-05
Year 82-92, n=240, $\rho=0.853915$, $a_{i1} \neq a_{i2}$		
5.5485177E-03	1.2031096E-05	4.2817146E-05
Year 82-92, n=216, $\rho=0.853915$, $a_{i1} = a_{i2}$		
2.6988420E-04	1.2610326E-05	1.2683163E-05

In these calculations variance of Y_i 's is assumed to be 1.045349 in all the years.

Now, we give the expressions for bias, variance and mean square error of $\hat{\beta}$ under AR(1) assumption by taking three measurements per subject.

In this case $\hat{\beta}$, $\text{var}(\hat{\beta})$ and bias are given by

$$\hat{\beta} = \frac{1}{2n\Delta} \sum (y_{i3} - y_{i1})$$

$$\text{var}(\hat{\beta}) = \frac{1}{2n\Delta^2} \sum \{\text{var}(y_{i3}) + \text{var}(y_{i1}) - 2\text{cov}(y_{i1}, y_{i3})\}.$$

and

$$\text{bias} = \frac{\tau}{\Delta} - \frac{1}{2n\Delta} \sum \{E(y_{i3}) - E(y_{i1})\}.$$

For convenience in notation, $Y_{i1} > a_{i1}$, $Y_{i2} > a_{i2}$, $Y_{i3} > a_{i3}$ is left out. Let X_{i1} , X_{i2} and X_{i3} be the standardized random variables corresponding to Y_{i1} , Y_{i2} and Y_{i3} respectively. Then the expected values of Y_i 's, variances of Y_i 's and covariance between Y_{i1} and Y_{i3} can be calculated using the following expressions that are derived from Tallis (1961).

$$E(X_{i1}) = \frac{1}{\alpha_i} \left[\phi(a_{i1}) \Phi^*(A_{i,12}, A_{i,13}; \frac{\rho}{\sqrt{1+\rho^2}}) + \rho \phi(a_{i2}) \Phi^*(A_{i,21}, A_{i,23}; 0) \right.$$

$$\left. + \rho^2 \phi(a_{i3}) \Phi^*(A_{i,31}, A_{i,32}; \frac{\rho}{\sqrt{1+\rho^2}}) \right]$$

$$E(X_{i2}) = \frac{1}{\alpha_i} \left[\rho \phi(a_{i1}) \Phi^*(A_{i,12}, A_{i,13}; \frac{\rho}{\sqrt{1+\rho^2}}) + \phi(a_{i2}) \Phi^*(A_{i,21}, A_{i,23}; 0) \right.$$

$$\left. + \rho \phi(a_{i3}) \Phi^*(A_{i,31}, A_{i,32}; \frac{\rho}{\sqrt{1+\rho^2}}) \right]$$

$$E(X_{i3}) = \frac{1}{\alpha_i} \left[\rho^2 \phi(a_{i1}) \Phi^*(A_{i,12}, A_{i,13}; \frac{\rho}{\sqrt{1+\rho^2}}) + \rho \phi(a_{i2}) \Phi^*(A_{i,21}, A_{i,23}; 0) \right.$$

$$\left. + \phi(a_{i3}) \Phi^*(A_{i,31}, A_{i,32}; \frac{\rho}{\sqrt{1+\rho^2}}) \right]$$

$$E(X_{i1}^2) = 1 + \frac{1}{\alpha_i} \left[a_{i1} \phi(a_{i1}) \Phi^*(A_{i,12}, A_{i,13}; \frac{\rho}{\sqrt{1+\rho^2}}) + \rho^2 a_{i2} \phi(a_{i2}) \Phi^*(A_{i,21}, A_{i,23}; 0) \right.$$

$$\left. + \rho^4 a_{i3} \phi(a_{i3}) \Phi^*(A_{i,31}, A_{i,32}; \frac{\rho}{\sqrt{1+\rho^2}}) + \rho(1-\rho^2) \phi(a_{i1}, a_{i2}) \Phi^*(A_{i,13}^2) \right]$$

$$+ \rho^2(1-\rho^2) \phi(a_{i1}, a_{i3}) \Phi^*(A_{i,12}^3) + \rho^2(1-\rho^3) \phi(a_{i2}, a_{i3}) \Phi^*(A_{i,21}^2) \]$$

$$E(X_{i3}^2) = 1 + \frac{1}{\alpha_i} \left[\rho^4 a_{i1} \phi(a_{i1}) \Phi^*(A_{i,12}, A_{i,13}; \frac{\rho}{\sqrt{1+\rho^2}}) \right.$$

$$\left. + \rho^2 a_{i2} \phi(a_{i2}) \Phi^*(A_{i,21}, A_{i,23}; 0) + \rho^4 a_{i3} \phi(a_{i3}) \Phi^*(A_{i,31}, A_{i,32}; \frac{\rho}{\sqrt{1+\rho^2}}) \right]$$

$$+ \rho^2 \phi(a_{i1}, a_{i2}) \Phi^*(A_{i,23}) + \rho^2(1-\rho^2) \phi(a_{i2}, a_{i3}) \Phi^*(A_{i,21}^3)$$

$$+ \phi(a_{i1}, a_{i3})(\rho-\rho^3) \Phi^*(A_{i,32}) + \phi(a_{i1}, a_{i3}) \rho^3(1-\rho^2) \Phi^*(A_{i,12}^3) \]$$

and

$$\begin{aligned}
E(X_{i1}X_{i3}) = & \rho^2 + \frac{1}{\alpha_i} [\rho^2 a_{i1} \phi(a_{i1}) \Phi^*(A_{i,12}, A_{i,13}; \frac{\rho}{\sqrt{1+\rho^2}}) \\
& + \rho^2 a_{i2} \phi(a_{i2}) \Phi^*(A_{i,21}, A_{i,23}; 0) + \rho^3 a_{i3} \phi(a_{i3}) \Phi^*(A_{i,31}, A_{i,32}; \frac{\rho}{\sqrt{1+\rho^2}}) \\
& + (\rho - \rho^3) \phi(a_{i1}, a_{i2}) \Phi^*(A_{i,23}) + (1 - \rho^4) \phi(a_{i1}, a_{i3}) \Phi^*(A_{i,32}) \\
& + \rho(1 - \rho^2) \phi(a_{i2}, a_{i3}) \Phi^*(A_{i,31}^2)],
\end{aligned}$$

where ϕ is the standard normal density function with $\Phi(a)$, the lower tail probability, and $\phi(x,y)$ is the standard bivariate normal density with correlation ρ and Φ^* is the corresponding upper tail probability,

$$\begin{aligned}
a_{i1} &= \frac{c_1 - y_{i1}^*}{\sigma}, \quad a_{i2} = \frac{c_2 - y_{i2}^*}{\sigma}, \quad a_{i3} = \frac{c_3 - y_{i3}^*}{\sigma}, \\
A_{i,12} &= \frac{a_{i2} - \rho a_{i1}}{\sqrt{1-\rho^2}}, \quad A_{i,21} = \frac{a_{i1} - \rho a_{i2}}{\sqrt{1-\rho^2}}, \\
A_{i,23} &= \frac{a_{i3} - \rho a_{i2}}{\sqrt{1-\rho^2}}, \quad A_{i,32} = \frac{a_{i2} - \rho a_{i3}}{\sqrt{1-\rho^2}}, \\
A_{i,13} &= \frac{a_{i3} - \rho^2 a_{i1}}{\sqrt{1-\rho^4}} \text{ and } A_{i,31} = \frac{a_{i1} - \rho^2 a_{i3}}{\sqrt{1-\rho^4}}.
\end{aligned}$$

Note that when the standardized cut points are same $A_{i,12} = A_{i,21} = A_{i,23} = A_{i,32}$ and $A_{i,13} = A_{i,31}$. In this special case bias is zero and $\hat{\beta}$ is an unbiased estimator of β . The variance of $\hat{\beta}$ is given by

$$\text{var}(\hat{\beta}) = \frac{\sigma^2}{25n^2} \{ \text{var}(X_{i1}) + \text{var}(X_{i3}) - 2\text{cov}(X_{i1}, X_{i3}) \},$$

where

$$\begin{aligned}
E(X_{i1}) = E(X_{i3}) &= \frac{(1+\sigma)}{\alpha_i} \phi(a_{i1}) \Phi(-A_{i,12}) \\
E(X_{i1}^2) &= 1 + \frac{1}{\alpha_i} [(1 + \rho^4) a_{i1} \phi(a_{i1}) \Phi^*(A_{i,12}, A_{i,13}; \frac{\rho}{\sqrt{1+\rho^2}}) \\
&+ \rho^2 a_{i1} \phi(a_{i1}) \Phi^*(A_{i,21}, A_{i,23}; 0) + \rho(1 - \rho^2) \phi(a_{i1}, a_{i1}) \{\Phi^*(A_{i,13}^2) \\
&+ \rho \Phi^*(A_{i,12}^2)\} + \rho^2(1 - \rho^3) \phi(a_{i1}, a_{i1}) \Phi^*(A_{i,21}^2)]
\end{aligned}$$

$$\begin{aligned}
E(X_{i3}^2) &= 1 + \frac{1}{\alpha_i} [2\rho^4 a_{i1} \phi(a_{i1}) \Phi^*(A_{i,12}, A_{i,13}; \frac{\rho}{\sqrt{1+\rho^2}}) \\
&+ \rho^2 a_{i1} \phi(a_{i1}) \Phi^*(A_{i,21}, A_{i,23}; 0) + \rho \phi(a_{i1}, a_{i1}) \{\rho \Phi^*(A_{i,23}^{-1}) \\
&+ \rho(1 - \rho^2) \Phi^*(A_{i,21}^3)\} + (1 - \rho^2) \phi(a_{i1}, a_{i1}) \Phi^*(A_{i,32}^{-1}) + \rho^2(1 - \rho^2) \Phi^*(A_{i,12}^3)\}]
\end{aligned}$$

and

$$\begin{aligned} E(X_{i1}X_{i3}) = & \rho^2 + \frac{1}{\alpha_i} [\rho^2(1+\rho)a_{i1}\phi(a_{i1})\Phi^*(A_{i,12}, A_{i,13}; \frac{\rho}{\sqrt{1+\rho^2}}) + \\ & + \rho^2 a_{i1}\phi(a_{i1})\Phi^*(A_{i,21}, A_{i,23}; 0) + \phi(a_{i1}, a_{i1})\{\rho(1-\rho^2)\Phi^*(A_{i,23}^{-1}) \\ & + (1-\rho^4)\Phi^*(A_{i,32}^{-1}) + \rho(1-\rho^2)\phi(a_{i1}, a_{i1})\Phi^*(A_{i,31}^{-2})\}]. \end{aligned}$$

Table 3.2

Year 82-87-92, n=240, $\rho=0.853915$, Standardized cuts Points are not same

bias	variance	MSE
1.2560963E-03	3.6896061E-05	3.8473839E-05
<hr/>		
4.3885500E-05	4.4242643E-05	4.4244569E-05

4. CONCLUSION

Half-life of 2,3,7, 8-tetrachlorodibenzo-p-dioxin (TCDD) in veterans of Operation Ranch Hand is studied using measurements from serum collected in 1982, 1987 and 1992. We estimate the half-life using two measurements per subject and three measurements per subject. These half-lives are corrected by finding the regression effect. Using 240 Ranch Hand veterans with dioxin measurements taken in 1982, 1987 and 1992 the corrected half-life is obtained as 7.77 years. For 1987, 1992 and 1982, 1987 dioxin measurements, the corrected half-lives are obtained as 8.6 years and 9.0 years respectively. Moreover, bias, variance and mean square error (MSE) of the estimators are computed and compared for various cases. In our study sample sizes are relatively large (over 200) and therefore the variances are very small. When we take two measurements per subject we see from table (3.1) that the mean square errors for the unbiased estimators are smaller than that of the biased estimators. On the other hand, when three measurements per subject are available biased estimator has a smaller mean square error than that of the

unbiased estimator. Since the MSE is very small for both cases we can still take the unbiased estimator. We observe that for the small sample size the same trend doesn't hold. Further more, we study the MSE as a function of the truncation point. We fix the second truncation point and move the the first one towards the second one. It is observed that the difference in mean square errors between unbiased and biased estimators changes sign. That is, when the truncation points are closed to each other unbiased estimator has a smaller mean square error than that of the biased estimator. When the first truncation point moves away from the second point, mean square error of the biased estimator is getting bigger than that of the unbiased estimator.

ACKNOWLEDGEMENTS

I would like to thank the Armstrong laboratory and the Air Force office of Scientific Research for sponsoring this research. This work was done under the guidance of Dr. Joel E. Michalek. The author thanks Dr. Michalek for his assistant in preparing this manuscript.

REFERENCES

James, K. E. (1973). Regression toward the mean in uncontrolled clinical studies. *Biometrics*, 29:121-130.

Michalek, J. E., Tripathi, R. C., Caudill, S. P. and Pirkle, J. L. (1992). Investigation of TCDD half-life heterogeneity in veterans of Operation Ranch hand. *J. Toxicol. Environ. Health* 35:29-38.

Pirkle, J. L., Wolfe, W. H., Patterson, D. G., Jr., Needham, L. L., Michalek, J. E., Miner, J. C. and Peterson, M. R. (1989). Estimates of the half-life of 2, 3, 7, 8-tetrachlorodibenzo-p-dioxin in Veterans of Operation Ranch Hand. *J. Toxicol. Environ. Health* 27:165-171.

Senn, S. S. and Brown, R. A. (1985). Estimating treatment effects in clinical trials subject to regression to the mean. *Biometrics*, 41:555-560.

Tallis, G. M. (1961). The moment-generating function of the truncated multinormal distribution. *J. R. Statist. Soc. B*, 23:223-229.

Wolfe, W. H., Michalek, J. E., Miner, J. C. Pirkle, J. L., Caudill, S. P., Patterson, Jr, D. G. and Needham, L. L. (1994). Determinants of TCDD half-life in veterans of Operation Ranch hand. *J. Toxicol. Environ. Health* 41:481-488.

Table 3.3

Year 82-87b, n=387, $\rho=0.88497$, $a_{i1} \neq a_{i2}$		
<u>bias</u>	<u>variance</u>	<u>MSE</u>
3.6093086E-03	2.3297681E-05	3.6324790E-05
Year 82-87, n=365, $\rho=0.88497$, $a_{i1} = a_{i2}$		
1.705659E-04	2.4414508E-05	2.4443601E-05
Year 87-92, n=240, $\rho=0.82286$, $a_{i1} \neq a_{i2}$		
1.0548597E-02	5.6190886E-05	1.6746379E-04
Year 87-92, n=213, $\rho=0.82286$, $a_{i1} = a_{i2}$		
3.6289920E-04	6.1123183E-05	6.1254879E-05
Year 82-92, n=240, $\rho=0.853915$, $a_{i1} \neq a_{i2}$		
5.5485177E-03	1.2031096E-05	4.2817146E-05
Year 82-92, n=216, $\rho=0.853915$, $a_{i1} = a_{i2}$		
2.6988420E-04	1.2610326E-05	1.2683163E-05
Year 82-87-92, n=240, $\rho=0.853915$, Standardized cuts are not same		
1.2560963E-03	3.6896061E-05	3.8473839E-05
Year 82-87-92, n=213, $\rho=0.853915$, $a_{i1} = a_{i2} = a_{i3}$		
4.3885500E-05	4.4242643E-05	4.4244569E-05

In these calculations variance of Y_i 's is assumed to be 1.045349 in all the years.

APPENDIX

Under AR(1) assumption, from Tallis (1961), we have for the bivariate density function

$$\begin{aligned}\alpha_i E(X_{i1}) &= \phi(a_{i1})\Phi^*(A_{i,12}) + \rho\phi(a_{i2})\Phi^*(A_{i,21}) \\ &= \phi(a_{i1})\Phi(-A_{i,12}) + \rho\phi(a_{i2})\Phi(-A_{i,21})\end{aligned}$$

$$\begin{aligned}\alpha_i E(X_{i2}) &= \rho\phi(a_{i1})\Phi^*(A_{i,12}) + \phi(a_{i2})\Phi^*(A_{i,21}) \\ &= \rho\phi(a_{i1})\Phi(-A_{i,12}) + \phi(a_{i2})\Phi(-A_{i,21})\end{aligned}$$

$$\begin{aligned}\alpha_i E(X_{i1}^2) &= \alpha_i + a_{i1}\phi(a_{i1})\Phi^*(A_{i,12}) + \rho^2 a_{i2}\phi(a_{i2})\Phi^*(A_{i,21}) + \rho(1-\rho^2)\phi(a_{i2}, a_{i1}) \\ &= \alpha_i + a_{i1}\phi(a_{i1})\Phi(-A_{i,12}) + \rho^2 a_{i2}\phi(a_{i2})\Phi(-A_{i,21}) + \rho(1-\rho^2)\phi(a_{i2}, a_{i1}; \rho)\end{aligned}$$

$$\begin{aligned}\alpha_i E(X_{i1}^2) &= \alpha_i + \rho^2 a_{i1}\phi(a_{i1})\Phi^*(A_{i,12}) + a_{i2}\phi(a_{i2})\Phi^*(A_{i,21}) + \rho(1-\rho^2)\phi(a_{i2}, a_{i1}) \\ &= \alpha_i + \rho^2 a_{i1}\phi(a_{i1})\Phi(-A_{i,12}) + a_{i2}\phi(a_{i2})\Phi(-A_{i,21}) + \rho(1-\rho^2)\phi(a_{i2}, a_{i1}; \rho)\end{aligned}$$

$$\begin{aligned}\alpha_i E(X_{i1}X_{i2}) &= \alpha_i + \rho a_{i1}\phi(a_{i1})\Phi(-A_{i,12}) + \rho a_{i2}\phi(a_{i2})\Phi(-A_{i,21}) \\ &\quad + (1-\rho^2)\phi(a_{i2}, a_{i1}; \rho),\end{aligned}$$

where

$$A_{i,12} = \frac{a_{i2} - \rho a_{i1}}{\sqrt{1-\rho^2}}$$

$$A_{i,21} = \frac{a_{i1} - \rho a_{i2}}{\sqrt{1-\rho^2}}$$

$$\alpha = P(X_{i1} > a_{i1}, X_{i2} > a_{i2})$$

and

$$\phi(a_{i2}, a_{i1}; \rho) = \frac{\phi(a_{i1})\phi(-A_{i,12})}{\sqrt{1-\rho^2}}$$

For trivariate normal density function, we have

$$\begin{aligned}\alpha_i E(X_{i1}) &= \phi(a_{i1})\Phi^*(A_{i,12}, A_{i,13}; \frac{\rho}{\sqrt{1+\rho^2}}) + \rho\phi(a_{i2})\Phi^*(A_{i,21}, A_{i,23}; 0) \\ &\quad + \rho^2\phi(a_{i3})\Phi^*(A_{i,31}, A_{i,32}; \frac{\rho}{\sqrt{1+\rho^2}})\end{aligned}$$

$$\alpha_i E(X_{i2}) = \rho \phi(a_{i1}) \Phi^*(A_{i,12}, A_{i,13}; \frac{\rho}{\sqrt{1+\rho^2}}) + \phi(a_{i2}) \Phi(A_{i,21}, A_{i,23}; 0)$$

$$+ \rho \phi(a_{i3}) \Phi^*(A_{i,31}, A_{i,32}; \frac{\rho}{\sqrt{1+\rho^2}})$$

$$\alpha_i E(X_{i,3}) = \rho^2 \phi(a_{i1}) \Phi^*(A_{i,12}, A_{i,13}; \frac{\rho}{\sqrt{1+\rho^2}}) + \rho \phi(a_{i2}) \Phi^*(A_{i,21}, A_{i,23}; 0)$$

$$+ \phi(a_{i3}) \Phi^*(A_{i,31}, A_{i,32}; \frac{\rho}{\sqrt{1+\rho^2}})$$

$$\alpha_i E(X_{i1}^2) = \alpha_i + a_{i1} \phi(a_{i1}) \Phi^*(A_{i,12}, A_{i,13}; \frac{\rho}{\sqrt{1+\rho^2}}) + \rho^2 a_{i2} \phi(a_{i2}) \Phi^*(A_{i,21}, A_{i,23}; 0)$$

$$+ \rho^4 a_{i3} \phi(a_{i3}) \Phi^*(A_{i,31}, A_{i,32}; \frac{\rho}{\sqrt{1+\rho^2}}) + \rho(1-\rho^2) \phi(a_{i1}, a_{i2}) \Phi^*(A_{i,13}^2)$$

$$+ \rho^2(1-\rho^2) \phi(a_{i1}, a_{i3}) \Phi^*(A_{i,2}^3) + \rho^2(1-\rho^3) \phi(a_{i2}, a_{i3}) \Phi^*(A_{i,21}^2)$$

$$\alpha_i E(X_{i3}^2) = \alpha_i + \rho^4 a_{i1} \phi(a_{i1}) \Phi^*(A_{i,12}, A_{i,13}; \frac{\rho}{\sqrt{1+\rho^2}})$$

$$+ \rho^2 a_{i2} \phi(a_{i2}) \Phi^*(A_{i,21}, A_{i,23}; 0) + \rho^4 a_{i3} \phi(a_{i3}) \Phi^*(A_{i,31}, A_{i,32}; \frac{\rho}{\sqrt{1+\rho^2}})$$

$$+ \rho^2 \phi(a_{i1}, a_{i2}) \Phi^*(A_{i,23}) + \rho^2(1-\rho^2) \phi(a_{i2}, a_{i3}) \Phi^*(A_{i,21}^3)$$

$$+ \phi(a_{i1}, a_{i3})(\rho-\rho^3) \Phi^*(A_{i,32}) + \phi(a_{i1}, a_{i3}) \rho^3(1-\rho^2) \Phi^*(A_{i,12}^3)$$

and

$$\alpha_i E(X_{i1} X_{i3}) = \rho^2 \alpha_i + \rho^2 a_{i1} \phi(a_{i1}) \Phi^*(A_{i,12}, A_{i,13}; \frac{\rho}{\sqrt{1+\rho^2}})$$

$$+ \rho^2 a_{i2} \phi(a_{i2}) \Phi^*(A_{i,21}, A_{i,23}; 0) + \rho^3 a_{i3} \phi(a_{i3}) \Phi^*(A_{i,31}, A_{i,32}; \frac{\rho}{\sqrt{1+\rho^2}})$$

$$+ (\rho-\rho^3) \phi(a_{i1}, a_{i2}) \Phi^*(A_{i,23}) + (1-\rho^4) \phi(a_{i1}, a_{i3}) \Phi^*(A_{i,32})$$

$$+ \rho(1-\rho^2) \phi(a_{i2}, a_{i3}) \Phi^*(A_{i,31}^2),$$

where $\phi(x,y)$ is the standard bivariate normal density with correlation ρ and Φ^* is the upper tail probability.

Chlorine Disinfection of Dental Unit Water Lines

**Larry R. Sherman
Professor
Department of Chemistry**

**University of Scranton
Scranton, PA 18510-4626**

**Final Report for:
Summer Faculty Research Program
Armstrong Laboratory**

**Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base
Washington, D.C.**

and

Armstrong Laboratory

September 1995

Chlorine Disinfection of Dental Unit Water Lines

Larry R. Sherman, Department of Chemistry, University of Scranton, Scranton PA 18510-4626

ABSTRACT

Potentially pathogenic microorganisms may participate in the formation of microbial biofilms on the interior walls of the plastic tubing used to supply coolant and irrigating water in dental units. Conscientious treatment procedures appear to be necessary to control the growth of dental unit waterline biofilms. There is little scientific basis for currently recommended treatment protocols for treatment and maintenance of dental water delivery systems. The summer research program was designed to test the effect of diluted commercial bleach (5.25% sodium hypochlorite) on biofilm colonizing dental unit water lines. A baseline bacteriological assay was performed on 24 dental units located at 4 separate USAF dental clinics located in the same metropolitan area. Twelve units at 3 clinics (designated Clinics I, II, and III) had been treated once weekly with 1:10 NaOCl for periods ranging from 6 months to 6 years. Twelve units at Clinic IV were selected for a 28 day study of the effectiveness of 2 dilutions (1:10, 1:100) of 5.25% sodium hypochlorite for control of biofilm growth in dental water lines. Eight units were treated weekly with either 1:10 or 1:100 solutions of commercial bleach. The 4 controls were treated only with sterile distilled water.

The units at Dental Clinics I and II had relatively low levels of colony forming units/mL (CFU/mL) ranging from 0 to 8.8×10^3 CFU/mL with a mean of 2.2×10^3 CFU/mL. The chlorine profile curves produced a distinct chlorine concentration (ppm) versus a cumulative volume profile. The chlorine concentration showed very little initial change, then rapidly fell to zero as the solutions were flushed from the units with distilled water. The total chlorine mass recovered at these clinics varied from 94-101.7% with the average being approximately 98%. Assays of planktonic bacteria at Clinics III and IV were much higher (Clinic III Range: 1×10^2 CFU/mL to 4.8×10^5 CFU/mL, Mean: 1.8×10^5 CFU/mL; Clinic IV Range: 5.6×10^4 CFU/mL to 1.1×10^6 CFU/mL, Mean: 4.7×10^5 CFU/mL) suggesting higher levels of biofilm contamination. The chlorine recovered from Clinics III and IV varied from 40-88% with the average being approximately 80%. The chlorine profile curves for the units at Clinics III and IV initially fell very rapidly, then more slowly throughout the profile study period. The difference in the profile curves may be due to absorption and reaction of the available chlorine by adherent biofilms. Subsequent assays of planktonic bacteria support this hypothesis.

At the conclusion of the 28 day treatment of Clinic IV units, the average chlorine recovery was 99.7% and the profile for a 1:10 sodium hypochlorite treatment yielded curves with the same characteristics as those at Clinics I and II, suggesting low levels of microbial contamination. Although each subsequent treatment of units with 1:100 NaOCl solution consumed less chlorine than the previous treatment, planktonic bacteria were still recovered at the end of the 28 day study (Range: 30 CFU/mL to 5×10^6 CFU/mL, Mean: 1.3×10^6). A plot of the percent chlorine recovered versus the log of the CFU yields inverse linear relationships. When linear regression of all samples (n=52) is performed using the percentage of active chlorine recovered versus log CFU/mL, an inverse relationship is revealed.

Introduction

Microbial biofilms are ubiquitous in nature and can be found virtually wherever there is moisture and a solid substrate¹. Although dental plaque that forms on teeth may well be the most studied biofilm, it is not the only one with which the dental profession should be concerned. Nearly as ubiquitous as plaque in dental offices, is the phenomenon of biofilm formation within the small bore plastic tubing used in dental units. These small conduits have been shown to harbor a remarkably diverse microflora including some opportunistic pathogens.²⁻⁵ Only two anecdotal reports of water line related disease have been reported. Martin described two *pseudomonas* wound infections in immuno-compromised patients,⁶ and Williams and Atlas described a fatal case of Legionellosis in a dentist that may have been related to contaminated water in his dental unit⁷. While Dental Unit Water Line (DUWL) contamination is not currently considered a major public health problem, there is reason for concern about the potential for infectious disease transmission from contaminated DUWLs.

Levels of bacterial contamination in DUWLs commonly exceed 100,000 Colony Forming Units/mL (10^5 CFU/mL)²⁻⁴. Currently there is no national standard for water used in routine dental treatment. While some infectious disease authorities feel that the proposed standard for US potable water supplies (500 CFU/mL) is the desirable level⁸, others feel that the ultimate goal should be zero CFU/mL⁹. Commercially available separate water reservoir systems (SWS) can be ordered as options on most new dental units or retrofitted to existing ones. These systems isolate the unit from municipal water and allow the use of water of known microbiological quality. They also permit introduction of cleaners and disinfectants into the systems. This study on the activity of chlorine and its effect on planktonic bacterial output in biofilm colonized dental units was undertaken to determine if the use of sodium hypochlorite can control or eliminate biofilms and keep dental coolant water free of planktonic bacteria.

Experimental

Equipment: Four A-DEC Dental Units (A-DEC Inc., Newburg, OR) at the Brooks Dental Clinic (Clinic I), and in 214B, Bldg. 125 (used for preliminary work only) Brooks AFB. Four A-DEC Dental Units at Kelly AFB (Clinic II). Four A-DEC Units at Randolph AFB (Clinic III). Twelve A-DEC Units at Lackland AFB (Clinic IV). Units at clinics I-III were equipped with separate water reservoir systems (SWS) (A-DEC Inc., Newburg, OR). Units at Clinic IV were initially connected to the municipal water supply, but were converted to separate reservoirs after baseline chemical and biological testing. SPC Samplers (Millipore, Corp., Bedford, MA) for quantitative recovery of heterotrophic mesophilic water bacteria.

Chemicals: 0.02 N sodium thiosulfate, primary grade iodine, potassium iodide, Clorox bleach (Purex Corporation, Oakland, CA) diluted with distilled water 1:100 (500 ppm active chlorine) and 1:10 (5000 ppm active chlorine). Concentrations (ppm) of free chlorine in the Clorox solutions were about 10% greater after dilution than would be expected based upon the analysis stated on the label.

Procedure: Chlorine assay: The thiosulfate was standardized against iodine using a known quantity of reagent grade iodine in KI and titrating to a blue starch-iodine end point. All bleach solutions used in the study were assayed by adding approximately 100 mg of KI to a known volume of solution (depending upon the concentration of chlorine), adding 1 mL of 5% sulfuric acid, and titrating with the standard thiosulfate to a blue starch-iodine end point. A description of the chemistry is give under the section Chemical Equations. The assayed bleach solutions used for treatment [Cl-0] were used as reference samples for calculating the percent recovery of chlorine.

Bacteriological assay: Pooled 100 mL water samples were collected from the unit's coolant water equipped handpiece and air water syringe lines. The samples consisted of approximately equal amounts of water from each line. Prior to sampling, handpieces and air water syringe tips were removed and the orifices were wiped with 70% isopropanol sponges to remove skin bacteria. SPC samplers were prepared according to manufacturer's instructions¹⁰ and incubated for 5 days at 25°C. Colonies were counted under a stereo microscope and results recorded as colony forming units per mL (CFU/mL). Identification of specific organisms was not attempted as part of this study.

Dental unit water line treatment and sample collection: Samples were obtained using three variations on the basic technique:

Basic sampling technique: Before treating each unit, water samples were collected for residual chlorine analysis. An additional pooled 100 mL water sample was drawn from the air water syringe and hand piece lines (without handpieces) for a bacteriological assay. Following baseline sampling, units at Clinic IV were disconnected from municipal water and equipped with SWS. The DUWLs were purged with air to remove as much residual water from the lines as possible, the reservoir was totally filled with either 1:10 or 1:100 bleach solution [sample Cl-0], and approximately 40 mL of bleach (equivalent to the volume of tubing in the unit) was loaded and flushed through the system [sample Cl-1]. After 10 minutes of contact time, the systems were air purged and the effluent collected [sample Cl-2]. The lines were then flushed with 500 mL of sterile deionized water (sdH₂O) and collected as 5 sequential samples of approximately 100 mL [samples Cl-3 to Cl-7]. The systems were air purged dry and allowed to remain dry overnight. The SWS was refilled with sdH₂O the next morning and a final 100 mL sample was collected [sample Cl-8]. Dental Clinic III had two SWS requiring larger volumes of hypochlorite to fill and flush [Cl-1 & Cl-2].

Technique 2: This technique was used for the samples collected at Dental Clinic IV on D+7 and D+14. The reservoirs were filled with a 1:10 bleach solution [sample Cl-0]. Approximately 40 mL of bleach was loaded using the SWS [sample Cl-1]. After 10 minutes, the system was purged of bleach and the purged solution was added to the previous sample [Cl-1]. The system was then flushed with 500 mL of sterile deionized water collected in a single container [Cl-2]. The system was purged with air and allowed to remain dry until the next day. No sample was collected the following morning.

Technique 3: This sampling technique was used only at Clinic IV on D+ 21 and D+28. The only variation on the basic technique was that the final sample [sample Cl-4] consisted of all remaining rinse water [Cl-4 through

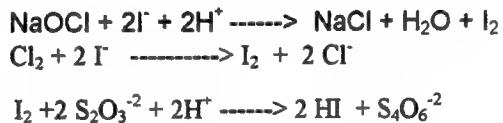
Technique 3: This sampling technique was used only at Clinic IV on D+ 21 and D+28. The only variation on the basic technique was that the final sample [sample Cl-4] consisted of all remaining rinse water [Cl-4 through Cl-7].

Table organization: All the tables and graphs are organized by dental clinics and by room number (e.g., LC217, in column B refers to the dental unit in C Wing, Room 217 of Clinic IV). Column C lists the active chlorine (ppm) in the bleach solutions [Cl-0] measured in the laboratory. The concentration of chlorine ejected from water lines when first loaded is found in column D [Cl-1]. Column E shows the concentration after 10 minutes of contact time and air purging [Cl-2]. The concentration of each effluent sample, when the system is flushed with sdH_2O , is found in columns F, G, H, I, and J [Cl-3 to Cl-7]. The concentration obtained the next morning is listed in column K [Cl-8]. The first row below the rooms lists the volume collected (mL) for each sample and the second row shows the cumulative volume.

The mass of chlorine was calculated by multiplying the total volume [Cl-1 + Cl-2] times the concentration in Cl-0. The amount of chlorine recovered was calculated by summing the product of effluent volume times the concentration for each sample ([Cl-1 through Cl-8]). The amount of chlorine consumed was the difference in the previous two calculations. The percentage of chlorine recovered was calculated from this difference.

Graphical representation: The profile curves of chlorine concentration (ppm) versus cumulative volume were plotted for the 1:10 bleach solutions at the four dental clinics. Graphs for the 1:10 and 1:100 bleach were prepared for Clinic IV on D+0, D+21, and D+28. D+7 and D+14 yielded only two points and therefore produced meaningless plots.

Chemical Equations:



5.25% NaOCl yields 5002 ppm active chlorine when quantitatively diluted 1:10.

1 ppm = 1 mg of Cl_2 per liter of solution or 1 μg of Cl_2 per mL.

$$N_{\text{S}_2\text{O}_3^{2-}} = (\text{mg of iodine}) / (126.905 \text{ mg/eq of I}_2 * \text{mL}_{\text{titrant}})$$

$$\text{Cl}_{\text{ppm}} = (N_{\text{S}_2\text{O}_3^{2-}} * 35.453 \text{ mg of Cl}_2/\text{eq} * \text{mL}_{\text{titrant}}) / (\text{aliquot})$$

Results

Pilot study: A pilot protocol was performed on one unit located in the research laboratory (Rm. 214B) and one unit located at Clinic I (Rm. B154). Since the unit in room 214B had not been treated or used in several months, it was reasonably assumed to be contaminated with biofilm. The unit in B154 had been treated once weekly with NaOCl for six years and had not shown evidence of extensive biofilm contamination on previous bacteriological assays. The two experimental treatments yielded different chlorine profile curves (99% chlorine was recovered for B154 versus 70% recovery for 214B); these preliminary findings suggested that chlorine analysis can be used to

measure the extent of microbial contamination in DUWL. The preliminary curves are not reproduced in this report because of their lack of precision. Another trial compared new, unused polyurethane water line tubing with identical tubing (from unit 214B) colonized with biofilm. A 12.15 cm piece of fresh tubing yielded a 96.2% recovery of the chlorine whereas a 14.2 cm piece of biofilm-contaminated tubing yielded only 80.8% recovery of the active chlorine. This second test gave further credence to the idea that chlorine recovery could be used as an indicator of microbial contamination in DUWL.

Background chlorine levels: Chlorine levels in all pre-treatment water samples--whether from contaminated dental units, tap water (collected from treatment room faucet), or water from the SWS--were below the detection limit (<0.2 ppm).

Dental Clinic I & II: The chlorine profile curves for Clinic I and Clinic II indicated that the chlorine concentration in the undiluted sample (collected by air purging the lines after the ten minute contact time) stays near the initial concentration. Flushing the lines with sdH_2O resulted in a rapid decrease in the chlorine content that can be observed in Tables B1 and K1. This type of profile curve was labeled a Type I Curve.

The total chlorine recovered varied from 352-470 mg and the percent recovery varied from 92.0 to 100.7% for the units in Clinics I and II. The average volume of chlorine solution was 75 mL for clinic I and 71 mL for clinic II. The average mass of chlorine consumed in the Clinic I & II study was 19.7 mg.

Dental Clinic III: Units at Dental Clinic III were equipped with assistant's carts which had a second SWS, increasing the total system volume to an average of 148 mL. This increased the average mass of chlorine from approximately 400 mg to 750 mg. The chlorine uptake varied from 55 to 310 mg and mean percent recovery of active chlorine was 80.4% (Range: 60.5% to 93.4%). The rapid decrease in chlorine concentration during flushing differs from the observations in Clinics I and II. Because of the large variation in the recovered chlorine, each set of chlorine data was individually graphed. A composite curve for all Clinic III units is presented in Table R1. The author labeled this type of curve a Type II Profile Curve. Correlation with planktonic bacterial counts support the hypothesis that Type II Profile Curves represent a set of microbially contaminated dental units¹⁰.

Dental Clinic IV: Twelve units at Clinic IV were used for a 28 day longitudinal study comparing the effect of two dilutions of 5.25% sodium hypochlorite (1:10 and 1:100) with sterile water controls. There are no chlorine data for the 4 control units since they were not treated with sodium hypochlorite. All these units were connected to municipal water supplies until immediately prior to initiation of the study. All 12 units initially demonstrated high levels of planktonic bacteria (Range: 5.6×10^4 to 1.1×10^6 CFU/mL, Mean: 4.7×10^5 CFU/mL). Planktonic bacterial counts for the 1:10 treatment group (n=4) ranged from 5.6×10^4 to 7.6×10^5 CFU/mL with a mean of 3.4×10^5 CFU/mL. The range for the 1:100 group (n=4) was 1.2×10^5 to 2.6×10^5 CFU/mL.

All 8 experimental units produced Type II Profile Curves as shown in Table L1. Although the average difference between samples [Cl-0] and [Cl-1] was small, there was a remarkable decrease in the amount of

All 8 experimental units produced Type II Profile Curves as shown in Table L1. Although the average difference between samples [Cl-0] and [Cl-1] was small, there was a remarkable decrease in the amount of available chlorine recovered between samples [Cl-1] and [Cl-2]. These points are much lower than the corresponding points found in a Type I Profile Curve. By the time 200 mL of bleach plus sterile water was flushed through the system, the chlorine concentration was small, but larger than in Clinics I & II. The average mass of chlorine consumed in the units on the 1:100 dilution was 12.2 mg whereas the quantity consumed in the 1:10 units was 48.9 mg. The mean percentage of chlorine recovered for the units subjected to 1:100 bleach was only 64.0%, whereas the average percent recovery from the units treated with 1:10 was 87%. Both of these values are below the 95.3% obtained with the units in Clinics I & II (Table L1).

Longitudinal study data:

After collection of baseline data, bacteriological and chlorine sampling was performed on the 8 experimental units at weekly intervals designated as D+7, D+14, D+21 and D+28. Only bacteriological data were collected on the control units. The following is a chronological report of results:

D+7: The quantity of chlorine consumed in the 1:100 treatment varied from 1 to 15 mg. The average amount of chlorine consumed was slightly less than at baseline (10.4 versus 12.2 mg). Likewise the average percentage recovered from the units treated with 1:100 bleach was greater (76.3% versus on 64 %). The average number of mg of Chlorine consumed in the 1:10 study was 51.8 mg versus 43.3 mg at baseline. The average amount of chlorine recovered was 89.1%, (Table L2). This is still smaller than the 95.3% recovered from the units at Clinics I and II.

D+14: The 1:10 treatment group consumed an average of 45.3 mg Cl chlorine (average Cl recovery: 89%). The 1:100 group consumed an average of 6.2 mg Cl (average Cl recovery: 84%). No curve was prepared because of the small number of points, (Table L3).

D+21: On D+16 technicians assigned to Suite C elected to treat all units, including the controls, with 1:10 NaOCl. Although this reduced planktonic colony counts, chlorine results for both 1:100 dilution units were consistent with the two 1:100 dilution units located in Suite A which were not treated (average mass of chlorine consumed in the two suites was 4.2 mg for Suite C and 4.5 mg for the units in Suite A). The percent of total chlorine recovered was approximately the same for both suites (90.4% and 91.7). A greater effect was evident with the units on the 1:10 dilution. The chlorine consumed in those units not subjected to the additional treatment varied between 38.8 and 42.7 mg, whereas those receiving the additional treatment varied from 2.7 to 6.0 mg. The percentage of chlorine recovered, however is much higher; 99.2% for the Suite C units compared to 91.5% for Suite A. The data is summarized in Table L3. The curve produced is a hybrid of Type I and Type II curves.

D+28: Both bacteriological and chemical analysis of samples collected on D+28 suggest a continuing decrease in adherent biofilm in dental water lines in sodium hypochlorite treated units when compared to control units which were treated only with sterile water. Planktonic bacterial counts in controls (n=4) ranged from 1.1 x

10^3 to 1.1×10^4 with a mean of 2.9×10^4 CFU/mL. In the 1:100 group, the average quantity of chlorine consumed fell to 3.8 mg and the quantity recovered increased to 94.7%. These results approach the 95.3% recovery characteristic of previously evaluated units with low levels of planktonic bacteria. Although the profile curves for the 1:100 group still exhibit a Type II configuration, they now more closely resemble the Type I curves associated with low planktonic bacterial counts. The planktonic bacterial counts for this group (Mean: 1.3×10^6) are dramatically skewed due to an anomalous result. SPC sample results for 3 of the 4 units ranged from 30 to 350 CFU/mL with a mean of 170 CFU/mL. Colony counts for the fourth unit (LA206) were too numerous to count at all dilutions, but were estimated at 5.0×10^6 CFU/mL. Despite such high planktonic counts from this unit, chlorine consumption (5.5 mg) and percent recovery (87.2%) were consistent with data for the other three units. This outlying data point is greater than 2 standard deviations from the mean and was not included in the linear regression analysis.

The average percent recovery for the NaOCl (1:10) group was 99.1% with a total chlorine consumption of only 7.0 mg of chlorine. All profile curves now exhibited a Type I configuration and very low levels of planktonic bacteria (Range 0 to 80 CFU/mL, Mean: 20 CFU/mL)¹⁰. The data and curves are found in Table L5.

Linear Regression Analysis:

Linear regression analysis of percent chlorine recovered versus log CFU/mL for all units (n= 52) revealed an inverse linear relationship between chlorine recovery and planktonic bacterial counts on SPC Samplers ($R^2 = 0.26$). The same procedure performed on combined data from Clinics I, II, and III (units previously treated with NaOCl) also showed an inverse relationship (Table Clinic I, II, III).

Similar results were obtained when linear regression was plotted for the 1:10 (n=31) and 1:100 (n=18) NaOCl treatment groups.

Discussion

Chlorine Profile Curves: This experiment produced two types of chlorine profile curve described as Type I and Type II. Type I curves are associated with units that have undergone repeated periodic treatment with diluted sodium hypochlorite, and exhibit low levels of planktonic bacteria in output water (< 100 CFU/mL). Type II curves result when sodium hypochlorite solutions are used to treat units which have high levels of planktonic bacteria (> 10^3 CFU/mL) in output water. There are two distinct components to these curves: (1) the treatment component represents a measurement of the loss of chlorine attributable primarily to reaction and absorption of chlorine during initial loading, contact time, and air purging of treatment solutions. (2). The dilution component represents the reduction in chlorine concentration resulting from dilution of the residual chlorine by the sterile distilled water used to flush the system after treatment..

Characteristic Type II curves exhibit an initial drop in chlorine concentration between the original treatment solution [Cl-0], the initial output sample [Cl-1] usually followed by a dramatic decrease in the concentration of

chlorine in the solution recovered by air purging the unit after 10 minutes of contact time [Cl-2] (Curve attached to Table L1). Although the initial decrease in chlorine content may be due to several factors including dilution, spillage and reaction with the biofilm, the decrease between [Cl-1] and [Cl-2] can best be explained only by absorption by and reaction with adherent biofilm within the dental waterlines. This loss of free chlorine is most likely due to the reaction of chlorine with organic molecules present in the biofilm glycocalyx. Although the mass of microbial biofilm adherent to the internal walls of water lines was not specifically measured, their presence and quantity can be inferred. Van der Wende *et al* have shown a mathematical relationship between planktonic cells in drinking water and biofilms¹².

By contrast, the relative lack of organic material available for reaction in treated units results in very little loss of chlorine content during treatment and produces a characteristic Type I curve. Most of the chlorine loss observed can be accounted for by dilution effects and loss of material due to minor spillage. For all practical purposes, the chlorine content remains constant (Curve attached to Tables B1 & K1).

After air purging the chlorine solutions from the water lines, they were flushed with approximately 500 mL of sterile distilled water to remove residual chlorine from the system. As with the loading component of the process, the flushing component curves demonstrate a distinctive geometry. Type I curves show a very precipitous drop in chlorine concentration as a result of the dilution process, approaching zero within the first 100 mL sampling interval. Type II curves, on the other hand, tend to exhibit a more gradual elimination of residual chlorine. This phenomenon may be due to the slow diffusion of chlorine from the biofilm matrix back into solution.

Due to the distribution of planktonic bacterial counts, there were insufficient data to generate chlorine profile curves for units with counts ranging between 10^2 and 10^3 CFU/mL. Since correlation analysis of all data points demonstrated an inverse linear relationship between numbers of planktonic bacteria and the percent of chlorine recovered, it is reasonable to assume that such curves will be intermediate in form.

Consumption and recovery of active chlorine : The percent of chlorine recovered following treatment appears to be related to the extent of microbial contamination in the system. When the quantity of chlorine recovered is greater than 95% the system shows little evidence of extensive biofilm contamination as measured by planktonic bacterial counts. All units which produced no growth on SPC Samplers¹⁰ exhibited Type I Profile Curves and chlorine recovery greater than 95%. When chlorine recovery was less than 90%, units all produced planktonic bacterial counts greater than 2×10^3 CFU/mL, suggestive of biofilm contamination.

Correlation analysis: Linear regression analysis of planktonic bacterial counts in comparison with the percentage of active chlorine recovered following sodium hypochlorite treatment revealed an inverse relationship. As the percentage of chlorine recovered approached 100% (reflecting decreased chlorine uptake within the system), the number of colonies recovered on SPC Samplers approached zero. Although this pattern held true for most data points there were "outliers". In one instance, bacterial counts from a single unit being treated with 1:100 NaOCl, which had previously exhibited a two log decrease in CFU/mL over a 3 week period, suddenly spiked colony

which had previously exhibited a two log decrease in CFU/mL over a 3 week period, suddenly spiked colony counts exceeding 5×10^6 CFU/mL. Despite this high level of bacterial growth, chlorine recovery data was almost identical to the other 3 units which exhibited an average of only 170 CFU/mL. This apparent anomaly may be explained by the observation that the 100 mL pooled water sample collected from that unit immediately prior to NaOCl treatment contained large amounts of macroscopic debris ejected from the handpiece lines and air water syringes. The sample taking apparently coincided with the break up of biofilm deposits, released large numbers of recoverable bacteria. Since the treatment regimen occurred subsequent to the biofilm breakup and clearance, chlorine recovery was not greatly affected.

The treatment protocol used in this study is similar to the one recommended by A-DEC and other manufacturers of dental units and after-market separate water reservoir systems¹⁴. This study suggests that conscientious compliance with a weekly treatment regimen can greatly reduce the levels of planktonic bacteria in dental unit output water over time. Use of a 500 mL flush following treatment results in insignificant quantities of residual chlorine in treatment water.

Theory: The mechanism for chlorine reaction with bacteria is still being debated (G. R. Dychdala, "Chlorine and chlorine Compounds" in *Disinfecting, Sterilization, and Preservation* ed. S. S. Block, Lea & Febiger, Philadelphia [1991])¹¹. There are questions whether the chlorine reacts with the protoplasm or other parts of the cell. When chlorine was added to the units contaminated with biofilms, macroscopic particulates, presumably biofilm glycocalyx, often flaked off. The iodine produced a blue coloration with the flakes suggestive of polysaccharides in the titration solutions. Previous studies have suggested that the complex fibrillar polysaccharides which comprise the glycocalyx protect biofilm bacteria from disinfectants including chlorines¹³. Better understanding of the profile curves during purging and flushing with sdH₂O may lead to an understanding of the process by which biofilm is removed. However, much more data is required before anything but speculation can be made.

Conclusions

A composite concentration versus volume profile curves for both the 1:10 and the 1:100 treatment shows a significant initial decrease in chlorine concentration when biofilm is suspected to be present in the dental unit. If no biofilm exists, the chlorine concentration remains near the 500 or 5000 ppm level until the system is flushed with sdH₂O. A plot of Log CFU/mL versus percent chlorine recovered yields an inverse relation indicating that a measure of the chlorine content in the effluent may yield a method for determining presence of biofilm in a DUWL. Once sufficient data is available, it is possible that the quantity of chlorine consumed during sodium hypochlorite treatment can be used as an indication of the quantity of biofilm in the DUWL.

Acknowledgments

The author wishes to acknowledgment technical assistance from S. E. Mills, DDS and T. J. Plamondon, DDS. He also wishes to thank Miss J. F. Nemeth whose organization skills and dedication to the project made it possible to finish during the RDL summer fellowship. He further wishes to thank the AF Office of Scientific Research for the financial support for this work.

GLOSSARY OF TERMS

purge—clearing DUWL with air

flush—passing water through the DUWL

dH₂O - distilled water

sdH₂O- sterile deionized water, USP Abbot Laboratories, North Chicago, IL

DUWL- dental unit water lines

DIS-Dental Investigation Services

Cl-x chlorine sample number

D+x day of treatment at Dental Clinic IV

References

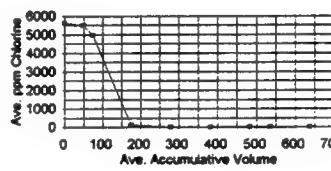
1. Costerton, J. W., Geesey, G. G., and Cheng, K. J. "How bacteria stick." *Sc. Am.* **238**:86-95 (1978).
2. Michell, R. and Borneff, M. "Significance of amoebae and other protozoa in water conduit systems of dental units." *ZBL. Bakt. Hyg.* **187**:312-323 (1989).
3. Schulze-Roebecke, R., Janning, B., and Fischeder. "Occurrence of mycobacterium in biofilm samples." *Tubercle & Lung Dis.* **73**:41-4 (1992).
4. Babcock, J. and Kleven, M. Bacteriological solution dispensers to eliminate bacterial contamination of spray units of dental chairs used in the VAC Dental Clinic, Wood, WI. Veterans Administration Clinic. (1977).
5. Williams, J. F., Johnston, A.M., Johnson, B., Huntington, M. K., and Mackenzie, C.D. "Microbial contamination of dental unit waterlines prevalence, intensity and microbiological characteristics." *JADA* **124**:59-65 (1993).
6. Martin, M. V. "The significance of bacterial contamination of dental unit water systems." *British Dent. J.* **163**:153-154 (1987).
7. Center for Disease Control. "Recommended Infection control practices for dentistry." *Morbidity and Mortality Weekly Reports.* **42**:Nr. Pr-8:7 (1993).
8. Safe drinking Water Act, Amended 1986; **42** USC, Public L. 99-339.
9. Levins, R. "Rethinking the cause of disease." *Chem. Eng. News* 37-8 (June 19, 1995).
10. Nemeth, J. F. "The analysis and disinfection of dental unit water lines". Final report. Research and Development Laboratory, Culver City CA (1995).
11. Dychdala, G. R. "Chlorine and chlorine compounds" in *Disinfecting, Sterilization, and Preservation* ed. S. S. Block, Lea & Febiger, Philadelphia (1991).
12. Van der Wende, E., W. G. Characklis & D. B. Smith "Biofilms and bacterial drinking water quality" *Water Res.* **23**:1313-1322 (1989).
13. Brown, M. R. W. & P. Giblet "Sensitivity of biofilms to antimicrobial agents". *J. Appl. Bacteriology* Supplement **74**:87s-97s (1993).
14. Mills S.E. in *Dental Items of Significance* 36 (37): 17 USAF Dental Investigation Service, Armstrong Laboratory, Brooks AFB TX (1992).

Chlorine Profile Curve B1 & K1

Dental Clinic I								Table B1								final95.wb2							
Chlorine Analysis ppm																							
sample	Cl-0	Cl-1	Cl-2	Cl-3	Cl-4	Cl-5	Cl-6	Cl-7	Cl-8	mg of Cl2	Recovered	%											
5 B183	5622.8	5877	4142		64.1	58.7	11.4	4.2	4.7	0 444.2	mg added												
6 Vol	0	51	28		94	99	105	98	93	99 429.59	mg recovered												
7 Acc Vol	0	51	79		173	272	377	475	568	667 14.612	mg consumed	96.7105											
8 B186	5622.8	5504	5436		132	2.48	1.33	1.58	1.77	0 466.69	mg added												
9 Vol	0	59	24		107	107	97	96	73	106 470	mg recovered												
10 Acc Vol	0	59	83		190	297	394	490	563	669 3.311	mg consumed	100.7095											
11 B158	5622.8	5337	5322		87.2	10.06	12.32	12.55	12.6	0 303.63	mg added												
12 Vol Col	0	32	22		105	106	101	107	45	100 301.24	mg recovered												
13 Acc Vol	0	32	54		159	265	366	473	518	618 2.384	mg consumed	99.21484											
14 B188	5622.8	5221	4895		144	7.05	0.35	0.29	0	0 461.07	mg added												
15 Vol Col	0	54	28		99	98	110	106	0	100 434.01	mg recovered												
16 Acc Vol	0	54	82		181	279	389	495	495	595 27.055	mg consumed	94.13199											
17 Ave ppm	5622.8	5484.8	4948.8		106.825	19.573	6.35	4.655	4.768	0 10.185	Ave. recovered	97.6917											
18 Ave Vol	0	49	74.5		175.75	278.25	381.5	483.25	536	637.25 11.696	Std. Dev.	2.503056											

Dental Clinic I

Chlorine Profile



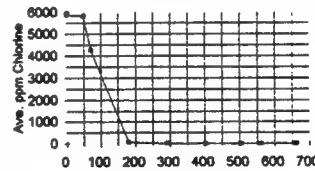
Dental Clinic II

Table K1

Dental Clinic II								Table K1								Chlorine Analysis ppm							
Chlorine Analysis ppm																							
sample	Cl-0	Cl-1	Cl-2	Cl-3	Cl-4	Cl-5	Cl-6	Cl-7	Cl-8	mg of Cl2	Recovered	%											
36 K124	5863.3	5813	3908.2		149.5	0.71	0.88	1.41	0.93	0 445.61	mg added												
37 Vol	0	56	20		93	96	95	102	106	99 417.99	mg recovered												
38 Acc Vol	0	56	76		169	265	380	462	568	667 27.621	mg consumed	93.80151											
39 K125	5863.3	5820.5	4515.3		38.8	0.58	0.55	0.19	4.45	0 381.11	mg added												
40 Vol	0	45	20		113	117	110	122	121	106 357.08	mg recovered												
41 Acc Vol	0	45	65		178	295	405	527	648	754 24.038	mg consumed	93.69281											
42 K126	5863.3	5782.6	4439.4		64.5	1.21	2.13	2.07	0	0 381.11	mg added												
43 Vol Col	0	42	23		113	119	121	99	0	100 352.87	mg recovered												
44 Acc Vol	0	42	65		178	297	418	517	517	617 28.244	mg consumed	92.58912											
45 K129	5863.3	5831.9	4213.4		49.07	0.696	0.327	4.51	0	0 463.2	mg added												
46 Vol Col	0	54	25		115	120	116	79	0	100 426.38	mg recovered												
47 Acc Vol	0	54	79		194	314	430	509	509	609 38.822	mg consumed	92.05046											
48 Ave Vol	0	49.25	71.25		179.75	292.75	403.25	503.75	560.5	661.75 29.181	Ave recovered	93.03348											
49 Ave ppm	5863.3	5812	4269.1		74.9675	0.799	0.9718	2.045	1.345	0 4.6946	Std. Dev.	0.739657											

Dental Clinic II

Chlorine Profile

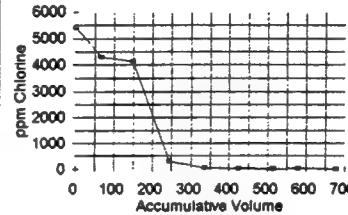


Chlorine Profile Curve R1 & L1

Dental Clinic III Table R1											Final 95b wb2			
2	Chlorine Analysis ppm													
3 B	C	D	E	F	G	H	I	J	K	L	M	N	%	
4 sample	Cl-0	Cl-1	Cl-2	Cl-3	Cl-4	Cl-5	Cl-6	Cl-7	Cl-8					
5 R118	5409	4454	3618	678.1	193.9	70.77	35.11	13.01	0	1001	mg added	Recovered		
6 Vol	0	79	106	129	122	103	96	54	99	857.9	mg recovered			
7 Acc Vol	0	79	185	314	436	539	635	689	788	142.7	mg consumed	85.7		
8 R212	5409	3767	2795	195.2	10.2	5.7	3.95	31.35	0	784.3	mg added			
9 Vol	0	47	98	99	98	96	94	64	106	474.2	mg recovered			
10 Acc Vol	0	47	145	244	342	438	532	596	702	310.1	mg consumed	60.5		
11 R214	5409	4972	4801	213.1	25.61	22.05	18.83	19.97	0	838.4	mg added			
12 Vol Col	0	53	102	101	97	87	99	91	100	782.9	mg recovered			
13 Acc Vol	0	53	155	256	353	440	539	630	730	55.51	mg consumed	93.4		
14 R219	5409	5367	4046	143.6	9.14	3.47	7.14	6.48	0 859	584.2	mg added			
15 Vol Col	0	83	25	56	48	72	70	49	87	479.6	mg recovered			
16 Acc Vol	0	83	108	164	212	284	354	403	490	104.5	mg consumed	82.1		
17 Ave Vol	0	65.5	148.3	244.5	335.8	425.3	515	579.5	677.5	153.2	Ave recovered	80.4		
18 Ave ppm	5409	4310	4145	307.5	59.71	25.5	16.26	17.7	0.215	95.69	Std. Dev.	12.2		

Dental Clinic III

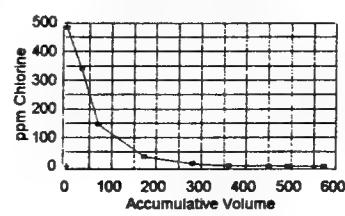
Chlorine Profile 1:10



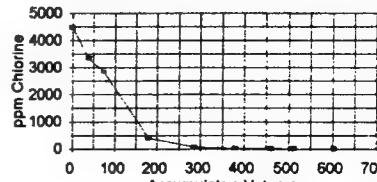
Dental Clinic IV Table L1
Chlorine Analysis ppm

1:100 dil										% Recovered	
	Cl-0	Cl-1	Cl-2	Cl-3	Cl-4	Cl-5	Cl-6	Cl-7	Cl-8		
LC217	483.5	361	237	33.8	5.9	2	1.5	0.82	0	28.53	mg added
Vol	0	19	40	104	109	108	116	48	20.75	mg recovered	
Acc Vol	0	19	59	163	272	380	496	544	7.772	mg consumed	72.8
LC218	483.5	260	146	12.6	3.6	1	0	0	0	32.88	mg added
Vol	0	38	30	106	102	100	60	0	80	16.06	mg recovered
Acc Vol	0	38	68	174	276	376	436	436	516	16.82	mg consumed
LA204	483.5	360	53	28.7	2.4	1.6	0	0	0	43.03	mg added
Vol	0	49	40	96	99	103	98	22	96	22.92	mg recovered
Acc Vol	0	49	89	185	284	387	485	507	603	20.11	mg consumed
LA206	483.5	377.7	145	61.7	25.8	4.2	2.3	0.1	0	29.98	mg added
Vol	0	27	35	109	121	122	90	30	96	25.84	mg recovered
Acc Vol	0	27	62	171	292	414	504	534	630	4.135	mg consumed
Ave Vol	0	33.25	69.5	173.3	281	362.3	451.3	493.3	573.3	12.21	Ave recovered
Ave ppm	483.5	339.7	145.3	34.2	9.425	2.2	0.95	0.23	0	6.492	Std Dev.
1:10 dil											
LA207	4470	3621	3048	322	56	13.5	11.9	0	0	366.5	mg added
Vol	0	40	42	102	103	101	38	0	99	313.3	mg recovered
Acc Vol	0	40	82	184	287	388	426	426	525	53.26	mg consumed
LA215	4470	3839	2855	354.6	58.8	16.5	8.9	0	0	366.5	mg added
Vol	0	43	39	109	106	112	66	0	95	323.7	mg recovered
Acc Vol	0	43	82	191	297	409	475	475	570	42.8	mg consumed
LC208	4470	2450	2149	670.8	83.4	39.5	2.72	0	0	219	mg added
Vol Col	0	26	23	100	100	50	100	100	100	190.8	mg recovered
Acc Vol	0	26	49	149	249	299	399	499	599	28.24	mg consumed
LC215	4470	3480	3328	256	56.5	16.4	8.8	6.8	2.64	371	mg added
Vol Col	0	45	38	112	116	114	117	104	72	322.1	mg recovered
Acc Vol	0	45	83	195	311	425	542	646	718	48.92	mg consumed
Ave Vol	0	38.5	74	179.8	286	380.3	460.5	511.5	603	43.3	Ave recovered
Ave ppm	4470	3348	2845	400.9	63.68	21.48	8.08	1.7	0.66	9.459	Std Dev.

Dental Clinic IV
Chlorine Profile 1:100



Dental Clinic IV
Chlorine Profile 1:10



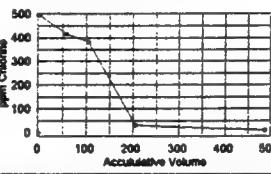
Chlorine Profile L2 & L3

Dental Clinic IV		Table L2											
2	Chlorine Analysis ppm												
3 B	C	D	E	F	G	H							
4 1:100 dilution						% Recovered							
5 sample	Cl-0	Cl-1	Cl-2										
6 LC217	496.2	385.3	26.4	56.96	mg added								
7 Vol	0	114.8	441	55.87	mg recovered								
8 Acc Vol	0	114.8	555.8	1.089	mg consumed	98.1							
9 LC218	496.2	365.85	5.27	66.79	mg added								
10 Vol	0	134.6	528.5	52.03	mg recovered								
11 Acc Vol	0	134.6	663.1	14.76	mg consumed	77.9							
12 LA204	496.2	386.37	9.45	68.23	mg added								
13 Vol	0	137.5	506.5	57.91	mg recovered								
14 Acc Vol	0	137.5	644	10.32	mg consumed	84.9							
15 LA206	496.2	341.7	10.3	64.06	mg added								
16 Vol	0	129.1	466	48.91	mg recovered								
17 Acc Vol	0	129.1	595.1	16.16	mg consumed	76.4							
18	Ave mg		10.33	Ave. recovered		84.3							
19 1:10 dilution	Std. Dev.		5.662	Std. Dev.		8.58							
20 LA207	5401	4648	128	677.8	mg added								
21 Vol	0	125.5	477.5	644.4	mg recovered								
22 Acc Vol	0	125.5	603	33.38	mg consumed	95.1							
23 LA215	5401	4495	137	793.4	mg added								
24 Vol	0	146.9	481.5	726.3	mg recovered								
25 Acc Vol	0	146.9	628.4	67.13	mg consumed	91.5							
26 LC208	5401	4696	119.8	747	mg added								
27 Vol Col	0	138.3	470.5	705.8	mg recovered								
28 Acc Vol	0	138.3	608.8	41.14	mg consumed	94.5							
29 LC215	5401	4604	95.2	755.1	mg added								
30 Vol Col	0	139.8	479.5	689.3	mg recovered								
31 Acc Vol	0	139.8	619.3	66.77	mg consumed	91.3							
32	Ave. Consumed		51.85	Ave. recovered		93.1							
33	Std. Dev.		40.6	Std. Dev.		1.7							
Dental Clinic IV		Table L3											
35	Chlorine Analysis ppm												
36 1:100 dilution						% Recovered							
37 sample	Cl-0	Cl-1	Cl-2										
38 LC217	525	361	15.18	29.4	mg added								
39 Vol	0	56	486	27.59	mg recovered								
40 Acc Vol	0	56	542	1.807	mg consumed	93.9							
41 LC218	525	365	12.7	45.15	mg added								
42 Vol	0	86	512.3	37.9	mg recovered								
43 Acc Vol	0	86	598.3	7.254	mg consumed	83.9							
44 LA204	525	340.6	9.46	61.43	mg added								
45 Vol	0	117	429	43.91	mg recovered								
46 Acc Vol	0	117	546	17.52	mg consumed	71.5							
47 LA206	525	380.2	12.38	41.48	mg added								
48 Vol	0	79	456	35.58	mg recovered								
49 Acc Vol	0	79	535	5.794	mg consumed	86							
50	Ave. Consumed		8.093	Ave. recovered		83.8							
51 1:10 dilution	Std. Dev.		5.795	Std. Dev.		8.03							
52 LA207	5924	4748	142	392.8	mg added								
53 Vol	0	66.3	485.5	383.7	mg recovered								
54 Acc Vol	0	66.3	551.8	9.028	mg consumed	97.7							
55 LA215	5924	3828	152.7	444.3	mg added								
56 Vol	0	75	415	350.5	mg recovered								
57 Acc Vol	0	75	490	93.83	mg consumed	78.9							
58 LC208	5924	4900.5	84.63	242.9	mg added								
59 Vol Col	0	41	372	232.4	mg recovered								
60 Acc Vol	0	41	413	10.48	mg consumed	95.7							
61 LC215	5924	4604	95.2	491.7	mg added								
62 Vol Col	0	83	436	423.6	mg recovered								
63 Acc Vol	0	83	519	68.05	mg consumed	86.2							
64	Ave. Consumed		45.35	Ave. recovered		89.6							
65	Std. Dev.		36.75	Std. Dev.		7.57							

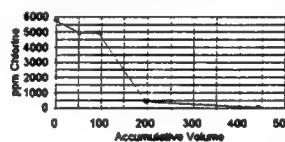
Chlorine Profile L4 & L5

1	Dental Clinic IV	Table L4			final95.wb2			
2		Chlorine Analysis ppm						
3	B	C	D	E	F	G	H	I
4	1:100 dilution							
5	sample	Cl-0	Cl-1	Cl-2	Cl-3	Cl-4		
6	LA204	495.6	397.2	345.5	273	7.93	40.6	mg added
7	Vol	0	40	42	99	404	36.3	mg recovered
8	Acc Vol	0	40	82	181	585	4.33	mg consumed
9	LA206	495.6	409	372.6	33.75	12.5	48.6	mg added
10	Vol	0	62	36	118	138	44.4	mg recovered
11	Acc Vol	0	62	98	216	352	4.12	mg consumed
12	LC217	495.6	406	390.8	29	7.6	55	mg added
13	Vol	0	43	68	90	332	49.2	mg recovered
14	Acc Vol	0	43	111	201	533	5.84	mg consumed
15	LC218	495.6	448.2	419.5	30.2	2.49	62.7	mg added
16	Vol	0	86	40.5	100	252	59	mg recovered
17	Acc Vol	0	86	126.5	226.5	478	3.68	mg consumed
18	Ave Conc.	495.6	414.6	342.1	30.08	7.82	4.49	Ave. recovered
19	Ave. Vol	0	57.75	104.4	206.1	487		
20		1:10 dilution			Std. Dev.	0.81	Std. Dev.	1.98
21	LA207	5756	4738	4443	445.2	35.1	478	mg added
22	Vol	0	40	43	108	318	439	mg recovered
23	Acc Vol	0	40	83	189	507	38.8	mg consumed
24	LA215	5756	4504	4799	431	31.1	489	mg added
25	Vol	0	41	44	98	300	446	mg recovered
26	Acc Vol	0	41	85	181	481	42.7	mg consumed
27	LC208	5756	5105	5301	385.4	37.2	530	mg added
28	Vol Col	0	56	36	108	247	527	mg recovered
29	Acc Vol	0	56	92	198	445	2.79	mg consumed
30	LC215	5756	5368	5249	375.3	41	668	mg added
31	Vol Col	0	69	47	107	110	862	mg recovered
32	Acc Vol	0	69	116	223	333	6.04	mg consumed
33	Ave ppm	5756	4928	4948	409.2	38.11	22.8	Ave recovered
34	Ave. Vol	0	51.5	94	197.8	4421	18.3	Std. Dev.
35								3.86

Dental Clinic IV L4
Chlorine Profile 1:100



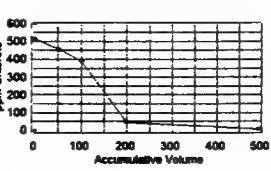
Dental Clinic IV L4
Chlorine Profile 1:10



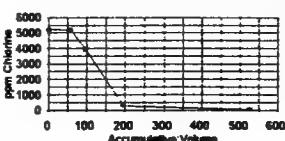
50 Dental Clinic IV Table L5 Chlorine Analysis ppm

51	1:100 dilution				%			
52	sample	Cl-0	Cl-1	Cl-2	Cl-3	Cl-4		Recovered
53	LA204	513.5	434.5	390.2	23.05	1.89	42.6	mg added
54	Vol	0	44	39	104	227	37.2	mg recovered
55	Acc Vol	0	44	83	187	414	5.48	mg consumed
56	LA208	513.5	471.1	379.2	36.46	3.55	58	mg added
57	Vol	0	40	73	103	386	51.7	mg recovered
58	Acc Vol	0	40	113	216	602	6.37	mg consumed
59	LC217	513.5	444.2	398.5	80.45	13.5	42.6	mg added
60	Vol	0	40	43	86	316	46	mg recovered
61	Acc Vol	0	40	83	189	485	3.4	mg consumed
62	LC218	513.5	469.5	396.4	40.92	4.28	65	mg added
63	Vol	0	88	40.5	100	252	61.6	mg recovered
64	Acc Vol	0	86	126.5	226.5	478	3.37	mg consumed
65	Ave Vol	0	52.5	101.4	199.6	495		
66	Ave ppm	513.5	454.8	390.8	45.22	5.82	Ave 1.295 mg	
67		1:10 dilution			Std. Dev.	3.82	Std. Dev.	8.13
68	LA207	5756	5039	4945	502.1	37.8	455	mg added
69	Vol	0	40	39	95	362	459	mg recovered
70	Acc Vol	0	40	79	174	536	4.1	mg consumed
71	LA215	5756	5121	5048	374.2	31.1	633	mg added
72	Vol	0	69	41	109	364	612	mg recovered
73	Acc Vol	0	69	110	219	583	20.8	mg consumed
74	Ave Conc.	5756	5033	2775	20	34.5	544	
75	Ave. Vol	0	54.5	94.5	196.5	500	9.84	
76	LC208	5756	5244	5107	398	26.1	696	mg added
77	Vol Col	0	84	37	100	279	676	mg recovered
78	Acc Vol	0	84	121	221	500	20.1	mg consumed
79	LC215	5756	5406	5028	375.3	66.6	478	mg added
80	Vol Col	0	40	43	86	318	489	mg recovered
81	Acc Vol	0	40	83	189	485	4.2	mg consumed
82	Ave Conc.	5756	5370	5068	385.6	46.4	587	
83	Ave. Vol	0	62	102	195	493	4.25	
84	Ave. ppm	5231	5202	3921	295.8	40.4	7.05	Ave. recovered 99.1
85	Ave. Vol	0	58.25	98.25	195.8	526	13.9	Std. Dev. 2.34

Dental Clinic IV L5
Chlorine Profile 1:100



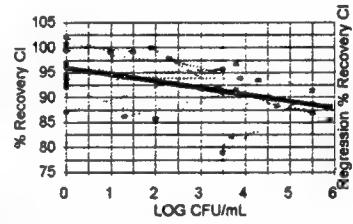
Dental Clinic IV L5
Chlorine Profile 1:10



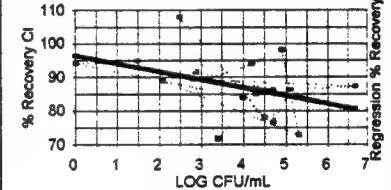
Composite Clinics I, II, III, & IV

26 Sept	5000ppm	Clinics I, II, III, & IV									
Room No.	LOG CFU	PRCT RECOVERY		Regress Values		BKR-LIN WB2					
B153	0	96.7	95.93495								
B155	0	100.7	95.93495								
B158	1.5	99.2	93.92817								
B166	1.5	94.1	93.92817	Constant	95.93495						
K124	3.9	93.8	90.71733	Std Err of Y Est	5.119439						
K125	0	93.7	95.93495	R Squared	0.215476						
K128	0	92.6	95.93495	No. of Observations	31						
K129	0	92	95.93495	Degrees of Freedom	29						
R118	2	85.7	93.25925								
R214	4.3	93.4	90.18219	X Coefficient(s)	-1.33785						
R216	3.7	82.1	90.9849	Std Err of Coef	0.474035						
LA207-1	5.9	85.5	88.04162								
LA207-2	3.3	95.1	91.52004								
LA207-3	2.3	97.7	92.85789								
LA207-4	3.4	91.9	91.38625								
LA207-5	1.9	100	93.39303								
LA215-1	4.7	88.3	89.64705								
LA215-2	3.8	91.5	90.85111								
LA215-3	3.5	78.9	91.25247								
LA215-4	5.5	91.3	88.57676								
LA215-5	3.8	96.7	90.65111								
LC208-1	0	87.1	95.93495								
LC208-2	0	94.5	95.93495								
LC208-3	3.5	95.7	91.25247								
LC208-4	0	99.5	95.93495								
LC208-5	0	97.1	95.93495								
LC215-1	5.5	86.8	88.57676								
LC215-2	3.8	91.3	90.85111								
LC215-3	1.3	86.2	94.19574								
LC215-4	1	99.1	94.59711								
LC215-5	0	102	95.93495	2.26129	92.90968	81.72901					
26 Sept				1.939684	5.590335						
Room No.				3.879367	11.18067						
LA204-2	4.3	84.9	86.09444								
LA203-3	3.4	71.6	88.24256								
LA204-4	2.8	89.3	89.67464	Constant	96.35767						
LA204-5	6.7	87.2	80.36613	Std Err of Y Est	8.517464						
LA206-1	5.1	86.2	84.185	R Squared	0.17298						
LA206-2	4.7	76.4	85.13972	No. of Observations	18						
LA206-3	4.7	86	85.13972	Degrees of Freedom	16						
LA206-4	2.9	91.5	89.43596								
LA206-5	2.1	89	91.3454	X Coefficient(s)	-2.3868						
LC217-1	5.3	72.8	83.70764	Std Err of Coef	1.304713						
LC217-2	4.9	98.1	84.66236								
LC217-3	4.2	93.9	86.33312								
LC217-4	3.2	89.4	88.71992								
LC217-5	2.5	108	90.39068								
LC218-2	4.5	77.9	85.61708	3.711111	87.5	69.50962					
LC218-3	4	83.9	86.81048	1.166308	8.995188						
LC218-4	0	94.1	96.35767	2.332615	17.99038						
LC218-5	1.5	94.8	92.77747								

1:10 Sodium Hypochlorite
Clinic I, II, III, & IV



1:100 Sodium Hypochlorite
Clinic IV



THE PROPER SEQUENCE FOR CORRECTING CORRELATION
COEFFICIENTS FOR RANGE RESTRICTION AND UNRELIABILITY

Joseph M. Stauffer
Department of Management & Finance

School of Business
Indiana State University
Terre Haute, IN 47809

Final Report for:
Summer Faculty Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Armstrong Laboratory

July 1995

THE PROPER SEQUENCE FOR CORRECTING CORRELATION
COEFFICIENTS FOR RANGE RESTRICTION AND UNRELIABILITY

Joseph M. Stauffer
Department of Management & Finance
Indiana State University

Abstract

Many researchers are unaware that the derivation for the formulas widely used in corrections for range restriction in correlation coefficients assume the absence of measurement error. Using the case where both restricted and unrestricted standard deviations are known for the variable upon which selection is made, it is shown that in order to apply the traditional range restriction correction formula correctly, the restricted correlation coefficient must first be corrected for unreliability in both variables. Additionally, the ratio of restricted to unrestricted standard deviations used in the correction must first be corrected for unreliability. Then, and only then, should the traditional formula be used. Similar results hold for other forms of range restriction. Implications for psychometric meta-analysis are briefly discussed.

THE PROPER SEQUENCE FOR CORRECTING CORRELATION COEFFICIENTS FOR RANGE RESTRICTION AND UNRELIABILITY

Joseph M. Stauffer

Basic formulas for the correction of correlation coefficients for range restriction are well known (Gulliksen, 1950; Thorndike, 1949). Perhaps the best known form of range restriction correction is the formula corresponding to what Thorndike (1949) labeled Case 2. Case 2 describes a situation where both the restricted and unrestricted standard deviations are known for the variable upon which direct selection occurs. Although found in a variety of algebraically equivalent forms, the correction for Case 2 is given by the formula

$$R_{xy} = r_{xy} / [u^2 (1 - r_{xy}^2) + r_{xy}^2]^{1/2}, \quad (1)$$

where R_{xy} is the unrestricted correlation between X and Y, r_{xy} is the restricted correlation, and u is the ratio of the restricted standard deviation of X to the unrestricted standard deviation of X, s_x/S_x .

What is apparently not well known is that the derivation for Equation 1 and other range restriction correction formulas make the assumption that the variables involved are measured perfectly reliably. This has serious implications for the wide-spread practice of correcting a correlation coefficient for both range restriction and unreliability.

This note derives the full range restriction correction formula for Case 2 and briefly discusses the theoretical implications of trying to correct the correlation of two unreliable

measures with a range restriction formula that assumes perfect reliability. Particular attention is paid to the impact on currently accepted psychometric meta-analysis methods.

Derivation of the Full Correction Formula

The derivation for Equation 1 begins by assuming that the relationship between two variables, say X^* and Y^* , is linear and that the slopes from the regression of Y^* on X^* in both the restricted and unrestricted samples are equal; that is,

$$B^* = b^*,$$

where B^* is the regression coefficient in the unrestricted sample and b^* is the regression coefficient in the restricted sample. Additionally, the conditional variances of Y^* given X^* are assumed equal in the unrestricted and restricted samples; that is,

$$S_{y^*}^2 (1 - R_{x^*y^*}^2) = s_{y^*}^2 (1 - r_{x^*y^*}^2),$$

where $S_{y^*}^2$ is the variance of Y^* in the unrestricted sample and $s_{y^*}^2$ is the variance of Y^* in the restricted sample. $R_{x^*y^*}$ and $r_{x^*y^*}$ are the correlations between X^* and Y^* in the unrestricted and restricted samples respectively.

Notice that these equalities rely on the assumption that X^* and Y^* are measured without error. When we consider the effects of measurement error, we have the fallible measures X and Y , and the above equality assumptions become

$$B/R_{xx} = b/r_{xx} \text{ and} \quad (2)$$

$$R_{yy} S_y^2 [1 - R_{xy}^2 / (R_{yy} R_{xx})] = r_{yy} s_y^2 [1 - r_{xy}^2 / (r_{yy} r_{xx})], \quad (3)$$

where B and b represent the slopes from the regression of Y on X in the unrestricted and restricted samples respectively, R_{xx} and r_{xx} are the corresponding reliabilities of X , R_{yy} and r_{yy} are the corresponding reliabilities of Y , and R_{xy} and r_{xy} are the correlations between X and Y in the unrestricted and restricted samples.

Converting Equation 2 from an expression of regression coefficients to that of correlation coefficients and squaring both sides for simplicity, we form the equality

$$[R_{xy}^2 / (R_{yy} R_{xx})] [(R_{yy} S_y^2) / (R_{xx} S_x^2)] = [r_{xy}^2 / (r_{yy} r_{xx})] [(r_{yy} s_y^2) / (r_{xx} s_x^2)].$$

Isolating $R_{xy}^2 / (R_{yy} R_{xx})$, the squared unrestricted and unattenuated correlation, we get

$$R_{xy}^2 / (R_{yy} R_{xx}) = [r_{xy}^2 / (r_{yy} r_{xx})] [(r_{yy} s_y^2) / (r_{xx} s_x^2)] [(R_{xx} S_x^2) / (R_{yy} S_y^2)].$$

This is equivalent to

$$R_{xy}^2 / (R_{yy} R_{xx}) = [r_{xy}^2 / (r_{yy} r_{xx})] [R_{xx} / (r_{xx} u^2)] [(r_{yy} s_y^2) / (R_{yy} S_y^2)],$$

with $u = s_x/S_x$. Substituting the result from Appendix A for the last term on the right-hand side of the equation yields

$$R_{xy}^2 / (R_{yy} R_{xx}) = \{ [r_{xy}^2 / (r_{yy} r_{xx})] [R_{xx} / (r_{xx} u^2)] \} / \{ 1 - [r_{xy}^2 / (r_{yy} r_{xx})] [1 + R_{xx} / (r_{xx} u^2)] \}.$$

Multiplying the right side of the equation by $[(r_{yy} r_{xx})/r_{xy}^2] / [(r_{yy} r_{xx})/r_{xy}^2]$ gives us

$$R_{xy}^2 / (R_{yy} R_{xx}) = [R_{xx} / (r_{xx} u^2)] / \{ [(r_{yy} r_{xx})/r_{xy}^2] - 1 + [R_{xx} / (r_{xx} u^2)] \}.$$

Further multiplying the right side by $[(r_{xx} u^2)/R_{xx}] / [(r_{xx} u^2)/R_{xx}]$ and taking the square root, we arrive at the full correction formula,

$$R_{xy} / (R_{yy} R_{xx})^{1/2} = \{ (r_{xx}/R_{xx}) u^2 [(r_{yy} r_{xx})/r_{xy}^2] - 1 \}^{-1/2}. \quad (4)$$

Equation 4 is useful for illustrative purposes (i.e., it shows that u and r_{xy} must be corrected for unreliability), but it fails to preserve the sign of r_{xy} . The computational formula to use for the correction should be

$$R_{xy} / (R_{yy} R_{xx})^{1/2} = r_{xy} / \{ (r_{xx}/R_{xx}) u^2 (r_{xx} r_{yy} - r_{xy}^2) + r_{xy}^2 \}^{1/2}. \quad (5)$$

It can be seen from Equations 4 and 5 that the full correction formula is identical to the traditional range restriction correction formula in Equation 1 when X and Y are free of

measurement error (i.e., $X = X^*$ and $Y = Y^*$). This means that the traditional range restriction correction formula is tenable if one first corrects for unreliability. As Equation 4 shows, this implies not only correction of the restricted correlation for both unreliability in X and Y using the restricted reliability estimates, r_{xx} and r_{yy} , but the square of the standard deviation ratio, u^2 , must be corrected for unreliability using the ratio of the restricted reliability of X , r_{xx} , over the unrestricted reliability of X , R_{xx} . Therefore, in order to apply the traditional formula for the correction of range restriction, one needs to estimate the restricted correlation, r_{xy} , the restricted reliability of X , r_{xx} , the unrestricted reliability of X , R_{xx} , the restricted reliability of Y , r_{yy} , and the standard deviation ratio, u .

In Case 2 the restricted correlation, r_{xy} , and standard deviation ratio, u , are taken as given. However, reliability estimates may be obtained from either the restricted sample or a sample regarded as unrestricted. Therefore, if unrestricted reliabilities are given, either R_{xx} or R_{yy} , estimates of the corresponding restricted reliabilities must be computed before applying the traditional correction for range restriction. If the restricted reliability of X is given, an estimate of the unrestricted reliability of X must be made.

To estimate R_{xx} given r_{xx} the formula found in Lord and Novick (1968, Equation 6.2.1, p. 130) is appropriate:

$$R_{xx} = 1 - [u^2 (1 - r_{xx})].$$

Conversely, the formula for estimating r_{xx} given R_{xx} is

$$r_{xx} = 1 - [(1/u^2)(1 - R_{xx})]. \quad (6)$$

If the ratio S_y/s_y is known, a formula analogous to Equation 6 estimates r_{yy} from R_{yy} :

$$r_{yy} = 1 - [(S_y/s_y)(1 - R_{yy})]. \quad (7)$$

If the ratio S_y/s_y is not known, the result from Appendix B is used:

$$r_{yy} = 1 - \{ [1 - (r_{xy}^2/r_{xx}) + (r_{xy}^2 R_{xx})/(r_{xx}^2 u^2)] [1 - R_{yy}] \}. \quad (8)$$

Implications for Meta-Analysis

This need to fully correct for unreliability before applying the range restriction correction formula has implications for psychometric meta-analysis, where corrections for range restriction are commonly made in conjunction with corrections for unreliability. Most available psychometric meta-analysis programs, for example, do not allow the distinction between restricted and unrestricted reliabilities to be made and, subsequently, fail to make the proper and necessary transformations. An exception that does allow the meta-analyst to designate reliabilities as either restricted or unrestricted, the BASIC program for individually corrected correlations presented in Hunter and Schmidt (1990), treats all reliabilities, for both X and Y, as the same (i.e., all are restricted or all are unrestricted). If unrestricted reliabilities are given, there is no attempt to convert them to restricted reliabilities, so that the correction for unreliability can

be made prior to correcting for range restriction. Also, no meta-analysis method either employs or suggests a correction to the standard deviation ratio, u , for unreliability.

Artifact distribution methods of psychometric meta-analysis are particularly unsuited for preserving the integrity of the correction sequence. Not only do these methods fail to allow the distinction between restricted and unrestricted reliabilities, thus making it possible to incorrectly form artifact distributions from a mix of restricted and unrestricted reliabilities, but they fail to properly consider the high degree of computational interdependence among the artifacts when estimating their means and variances. For example, to correct the observed restricted mean correlation for unreliability and range restriction, we must first estimate the mean r_{yy} , mean r_{xx} , mean R_{xx} , and mean u . Unless all reliability estimates for Y are restricted or an estimate of the mean of the ratio S_y/s_y is available, Equation 8 shows that the mean r_{yy} cannot be estimated without first estimating the mean r_{xx} , mean R_{xx} , and mean u . Furthermore, we must have an estimate of r_{xx} for every R_{xx} , and vice versa, in order to estimate both the mean r_{xx} and the mean R_{xx} . We cannot do any of this until we have an estimate for the mean u . This dictates a strict sequence for the artifact distribution analysis: estimate the mean u , estimate r_{xx} for each R_{xx} given, estimate R_{xx} for each r_{xx} given, estimate mean r_{xx} and mean R_{xx} , estimate r_{yy} for each R_{yy} , estimate the mean r_{yy} , and, finally, correct the mean r_{xy} .

Conclusion

Even though this note focused on one particular form of range restriction, the same principle applies to correction formulas for all basic forms (e.g., Equations 5, 7, and 8, Thorndike, 1949, pp. 173-175): Because these correction formulas were derived under the assumption of no measurement error, one must correct all variances and functions of variances

(e.g., correlation coefficients and standard deviations) for unreliability before applying the range restriction correction. This suggests the need to reassess meta-analytic methods for correcting correlation coefficients, which fail to recognize this important principle.

References

Gulliksen, H. (1950). Theory of mental tests. New York: Wiley.

Hunter, J. E., & Schmidt, F. L. (1990). Methods of meta-analysis: Correcting error and bias in research findings. Newbury Park, CA: Sage.

Lord, F. M., & Novick, M. R. (1968). Statistical theories of mental test scores. Menlo Park, CA: Addison-Wesley.

Thorndike, R. L. (1949). Personnel selection: Test and measurement techniques. New York: Wiley.

Appendix A

This appendix derives the formula for $R_{yy} S_y^2$. From Equation 3 we have

$$R_{yy} S_y^2 - [R_{yy} S_y^2 R_{xy}^2 / (R_{xx} R_{yy})] = r_{yy} s_y^2 - [r_{yy} s_y^2 r_{xy}^2 / (r_{xx} r_{yy})].$$

Isolating $R_{yy} S_y^2$ and translating the correlation coefficients to regression coefficients yields

$$R_{yy} S_y^2 = r_{yy} s_y^2 - (b/r_{xx})^2 s_x^2 + (B/R_{xx})^2 S_x^2.$$

Due to Equation 2,

$$R_{yy} S_y^2 = r_{yy} s_y^2 - (b/r_{xx})^2 s_x^2 + (b/r_{xx})^2 S_x^2.$$

Finally, converting the regression coefficients back into correlation coefficients and simplifying gives us

$$R_{yy} S_y^2 = r_{yy} s_y^2 \{ 1 - [r_{xy}^2 / (r_{xx} r_{yy})] [1 + (R_{xx}/r_{xx})(1/u^2)] \}. \quad (A1)$$

Appendix B

This appendix derives the formula for computing the restricted reliability for Y from its unrestricted reliability when the ratio S_y/s_y is not known.

From Equation A1 in Appendix A we find that

$$r_{yy} s_y^2 \{ 1 - [r_{xy}^2 / (r_{xx} r_{yy})] [1 + (R_{xx}/r_{xx})(1/u^2)] \} = R_{yy} S_y^2.$$

Isolating r_{yy} and simplifying gives us:

$$r_{yy} = R_{yy} (S_y^2 / s_y^2) + \{ (r_{xy}^2 / r_{xx}) [1 + R_{xx} / (r_{xx} u^2)] \}.$$

From Equation 7,

$$1 - (S_y^2 / s_y^2) + R_{yy} (S_y^2 / s_y^2) = R_{yy} (S_y^2 / s_y^2) + \{ (r_{xy}^2 / r_{xx}) [1 + R_{xx} / (r_{xx} u^2)] \}.$$

Solving for S_y^2 / s_y^2 and simplifying results in the equation

$$S_y^2 / s_y^2 = 1 - (r_{xy}^2 / r_{xx}) + (r_{xy}^2 R_{xx}) / (r_{xx}^2 u^2).$$

Finally, substituting this result for S_y^2 / s_y^2 in Equation 7 gives us

$$r_{yy} = 1 - \{ [1 - (r_{xy}^2 / r_{xx}) + (r_{xy}^2 R_{xx}) / (r_{xx}^2 u^2)] [1 - R_{yy}] \}. \quad (B1)$$

**Pharmacological Intervention to Increase the Time before Gravity Induced
Loss of Consciousness in Rats and Mice**

**William B. Stavinoha, PhD
Professor in Pharmacology
Department of Pharmacology**

**7703, Floyd Curl Drive
UT Health Science Center
San Antonio, Tx 78284**

**Final Report for:
Summer Faculty Research Program
Armstrong Laboratory**

**Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC**

and

Armstrong Laboratory

September 1995

Pharmacological Intervention to Increase the Time before Gravity Induced Loss of Consciousness in Rats and Mice

William B. Stavinoha, PhD
Professor in Pharmacology
Department of Pharmacology

Abstract

Research on the modification of +Gz induced loss of consciousness in rats and mice focused on the effect of +Gz on regional brain neurochemistry and the testing of cholinergic agents and an adenosine receptor blocking agent on the time to GLOC. Exposure of mice to +35Gz for 30 seconds produced GLOC. Thirty seconds of +35Gz significantly increased lactate in the cerebral cortex, cerebellum, hippocampus, midbrain and brain stem. It caused no change in acetylcholine but significant increases in choline in all regions. Significant decreases occurred in 5 hydroxytryptamine in cerebellum, brain stem and midbrain. Homovanillic acid increased significantly in the midbrain. Epinephrine decreased significantly in the cerebellum and mid brain. Chronic caffeine ingestion for 8 weeks decreased the dopamine in the cerebellum and brainstem in control mice and in GLOC mice. In the study of GLOC modification one cholinergic agent was effective in extending the time to GLOC in rats. 3,4-diaminopyridine 2.5 pmol/kg ip increased the time to GLOC from control times of 13.5 ± 2.7 to 24.7 ± 7.2 seconds while 5 pmol/kg increased the time to 81.7 ± 15.3 seconds. Caffeine an adenosine A1 receptor blocker at 50mg/kg increased the time to GLOC from 14.5 ± 2.3 seconds to 35 ± 13.2 seconds. These initial studies indicate that it is possible to significantly extend the time to GLOC by pharmacological intervention.

Pharmacological Intervention to Increase the Time before Gravity Induced Loss of Consciousness in Rats and Mice

William B. Stavinoha, PhD

Introduction

Pilots flying high performance aircraft are exposed to high acceleration forces. When a circular pattern acceleration causes a head to foot inertial load(high+Gz) blood is forced from the head to the body. This loss of blood from the brain, ischemia, can cause +Gz induced loss of consciousness (GLOC). GLOC is a significant concern both in terms of pilot safety and the limits its occurrence places on aircraft performance. A physical and physiological approach to the problem is the use of a G suit which inflates to increase pressure on the body to limit blood pooling. This action provides some extension of time to GLOC (Forster et al., 1994). The experimental use of glucose in rodents to extend the time to GLOC(Shahed et al., 1995) indicates that replacement therapy can be effective. An obvious extension of this research is the study of pharmacological intervention in GLOC. The Armstrong Laboratory has developed a number of unique approaches that lessen or eliminate some of the problems inherent in the study of GLOC. The studies reported here are based on the utilization of these unique resources. The basic research on GLOC has reached a level of information and techniques from which it is now possible to attempt a reasoned pharmacological intervention to extend the time to GLOC or diminish its duration and after effects. This project utilizes a pharmacological approach.

Discussion of the problem

Acceleration GLOC is caused by the loss of blood from the brain but the study of GLOC is distinguished by its transient nature from the many studies of ischemia in the literature. These characteristics engender several special approaches to the problem of GLOC. Special techniques must be utilized so as to provide information on the neurochemical status of brain regions during the short time interval of GLOC. The literature of ischemia must be interpreted with caution because the long ischemic interval utilized probably has little relation to GLOC. The studies have also been performed with many potentially confounding stresses such as strict confinement, surgery and anesthesia which are best avoided if possible. An approach suggested by the dream-like events that human subjects have reported during GLOC (Forster and Whinnery, 1988) is the study of the neurochemistry and pharmacological interventions suggested in the extensive literature on

sleep.

The approaches to the problem of studying GLOC and pharmacological intervention for GLOC were based on: (1) providing a secure space for the mouse that produced minimal stress both before and during high gravity studies (2) utilizing the Armstrong Laboratories system and techniques for monitoring consciousness level in animals to study the effect of pharmacological intervention on GLOC (3) evaluating the effects of GLOC on regional neurochemistry of the mouse thereby providing information for specific drug intervention.

1. A system of low stress mouse containment is important, for stress and prior level of metabolic activity of the brain governs the time that membranes depolarize in response to ischemia (Astrup et al., 1980). Also, the mouse must remain quite and in place during high +Gz to maintain the original tuning configuration and calibration of the microwave instrument.
2. The Armstrong Laboratory system of monitoring the EEG to monitor consciousness provides the means to identify the EEG isoelectric point during Gz studies and can be utilized as the point of GLOC. The time GLOC occurs is consistent both in the same animal and between animals in the same group. This ability to replicate GLOC times provides a powerful method to study the effects of pharmacological intervention on GLOC. The choice of drugs to study was based on the literature of ischemia and sleep. The acceleration produced ischemia reduces the oxygen supply of the brain and switches the metabolism from aerobic to anaerobic resulting in a great decrease in efficiency. It is known that small reductions in energy charge increase synthesis of adenosine from AMP. This can occur both in early ischemia (Winn et al., 1979) and as a manifestation of increased sleep need.(Bennington and Heller, 1995) Adenosine release will diminish the release of neurotransmitters such as acetylcholine (Phillis and Wu, 1981; Prince and Stevens, 1992). Neurons releasing acetylcholine are more active during waking than during sleep (Steriade et al., 1990). Caffeine is an adenosine A1 receptor blocking agent (Fredholm, 1995) Chronic caffeine ingestion has been reported to protect against ischemia by upregulating the A1 receptors (Rudolphi et al., 1989) Therefore both acute and chronic caffeine were studied. The low oxygen that occurs during early or mild ischemia diminishes mental acuity (Freeman et al., 1987). Also impaired acetylcholine synthesis accompanies mild hypoxia and hypoglycemia in mice (Gibson and Blass, 1976). Cholinergic agents such as 3,4 diaminopyridine protect against the deleterious effects of hypoxia (Peterson and Gibson, 1982). Therefore a series of cholinergic agonists and antagonists were studied for their ability to modify GLOC .
3. Research on the effects of acceleration and GLOC on regional levels of neurotransmitters was

initiated to provide basic information on the regional changes. The unconsciousness of sleep seems to strongly involve the thalamo-cortical axis (Steriade, 1993) and initial changes may occur here during the early ischemic GLOC.

Methodology

The design and fabrication of the constraint housing for the mouse utilized dental techniques to cast an exact replica of a mouse head. This head was then used to cast an acrylic helmet. The helmet was constructed in two halves to allow it to be fitted comfortably. Two acrylic guide rods assured a matching fit of the two halves. This helmet was then fitted onto a guide to maintain exact orientation for the head.

For monitoring of the EEG, electrodes were surgically placed in the skull using sterile techniques and halothane as anesthetic. After recovery the implanted mouse mouse or rat was placed in the centrifuge and connected to the EEG. The isoelectric point of the EEG was utilized as the time of GLOC during the +Gz tests. The drugs were given ip fifteen minutes before the +Gz run. The rats were exposed to +22.5Gz and the mice to +35Gz. Each animal served as its own control and was studied in several +Gz tests. A small number of +Gz runs does not change the animals GLOC time. In the study of chronic caffeine ingestion 0.2 % caffeine was placed in the drinking water of mice for 8 weeks. All the drugs used were obtained from Sigma.

The studies of regional brain neurochemistry was accomplished in mice sacrificed in the microwave centrifuge. The microwave operated at 4.3 Kw for 720 mseconds(Stavinoha, 1993) The catecholamines were assayed using an HPLC equipped with an electrochemical detector(Liptrot et al., 1993). Acetylcholine and choline was also assayed by HPLC (Bioanalytical Systems Kit MF 8910 West Lafayette Indiana). Lactate was assayed as previously reported.(Shahed et al., 1994).

Results

As shown in Table 1 acute caffeine given to rats at 30 or 50mg/kg ip increased time to GLOC. The higher dose of caffeine was less effective. There were no remarkable changes in recovery time. In mice 50 mg/kg ip increased time to GLOC, but chronic ingestion of caffeine had no notable efect.(Table 2) Cholinergic agents such as 3,4 diaminopyridine at a dose of 2.5pmol/kg ip extended the time to GLOC from 13.5 ± 2.7 to 24.7 ± 7.2 seconds and a dose of 5.0pmol/kg ip

increased the time to GLOC to 81.7 ± 15.3 seconds. One of the actions of 3,4 diaminopyridine is to increase the release of acetylcholine from nerve endings (Peterson and Gibson, 1982) thus possibly counteracting the adenosine inhibition of acetylcholine release. Physostigmine salicylate at a dose of 0.25mg/kg ip or 0.5mg/kg did not change the time to GLOC but did reduce the recovery time from GLOC from 32.5 ± 3.5 to 15.0 or with the larger dose to 20.0 seconds. Physostigmine is a reversible inhibitor of acetylcholinesterase the enzyme responsible for terminating the action of acetylcholine. Its presence increases the level of acetylcholine because of less destruction of the released acetylcholine. In GLOC it may mitigate the decreased release of acetylcholine due to adenosine release (Table 3).

EEG records were not obtained from mice used for regional metabolism studies because the EEG electrodes can interfere with the microwave field(Stavinoha, 1980). EEG instrumented mice were studied for time of GLOC and it was found that +35Gz for 30seconds resulted in GLOC. Therefore this was the time and speed used for subsequent studies of regional brain neurochemistry.

The catecholamines and metabolites norepinephrine, epinephrine, dopamine, homovanillic acid, 5 hydroxytryptamine, 3,4-dihydroxyphenyl acetic acid and 5-hydroxy indole acetic acid were assayed in mouse brain regions following +35Gz for 30 seconds in control and caffeine treated animals. Significant decreases in 5HT occurred in the cerebellum, brain stem and midbrain. Epinephrine showed significant decreases in the cerebellum and brain stem. A significant increase in HVA was noted in the midbrain.(Table 4) (Fig 3a and 3b)

Lactate, a very sensitive marker for ischemia, was measured in mouse brain regions. Lactate was increased significantly in all the brain regions studied following +35Gz for 30 seconds (Fig 1). No change occurred in acetylcholine in any region studied after +35Gz for 30 seconds. Choline increased significantly in all regions. Choline accumulation appears to be a sensitive marker compound for ischemia. Production of free choline is a rapid process (Stavinoha and Weintraub 1976) and does not require energy.

Conclusions

The use of a tailored helmet for the mouse improves fit and comfort and allows a more precise orientation of the mouse for the microwave and centrifuge studies. Use of the microwave centrifuge instrumentation provides the means to identify changes in the neurochemistry of the brain of the mouse during the GLOC process. This information should enhance the process of

pharmacological intervention both by improving knowledge of the neurochemical changes that occur and the effects of drugs on the processes . EEG monitoring of the test animals to identify GLOC time has proven to be a very useful technique for the study of pharmacological intervention in the process of GLOC. This essentially preliminary study indicates that both cholinergic agents and an adenosine receptor blocking agent, caffeine, can significantly modify GLOC. Caffeine is a safe drug in common use. . The possibility for success with drug intervention appears quite positive. The extension of time to GLOC by only a few seconds could be very significant in the protection of pilots. This approach to modification of GLOC has a great potential to discover agents to modify GLOC.

Table 1
Effect of Caffeine on GLOC in Rats

Treatment	Dose mg/kg	No. of Tests	Time in sec.	mean± sd
			GLOC	Recovery
None	0	8	14.5± 2.3	33.1±17.9
Caffeine	30	2	21.0±12.7	23.5± 9.2
Caffeine	50	3	35.0±13.2	56.3±52.4
Caffeine	70	2	17.5± 3.5	22.5± 3.5

Table 2
Effect of Caffeine on GLOC in mice

Treatment	Dose	No. of Tests	Time in sec.	mean \pm sd
			GLOC	Recovery
Control	0	4	18.3 \pm 3.1	44.5 \pm 5.3
Chronic Caffeine	0.2%	4	22.8 \pm 9.1	38.0 \pm 6.6
Acute Caffeine	50mg/kg	4	26.5 \pm 3.3	40.0 \pm 10.0

Table 3
Effect of Cholinergic agents in GLOC in Rats

Treatment	Dose	No. of Tests	Time in sec.	mean\pmsd
			GLOC	Recovery
Control	0	8	13.5 \pm 2.7	31.6 \pm 19.9
3,4-Diaminopyridine	2.5pmol/kg	3	24.7 \pm 7.2	43.7 \pm 23.3
3,4-Diaminopyridine	5.0pmol/kg	3	81.7 \pm 15.3	45.0 \pm 21.2
Control	0	2	15.0 \pm 0.0	32.5 \pm 3.5
Physostigmine	0.50mg/kg	1	15.0 \pm 0.0	20.0 \pm 0.0
Physostigmine	0.25mg/kg	1	15.0 \pm 0.0	15.0 \pm 0.0
Control	0	9	11.3 \pm 1.8	45.6 \pm 24.6
Atropine sulphate	16mg/kg	2	10.0 \pm 0.0	75.0 \pm 21.2
Atropine sulphate	25mg/kg	2	15.0 \pm 0.0	32.5 \pm 17.7
Atropine sulphate	100mg/kg	2	11.0 \pm 1.4	57.5 \pm 38.9
Atropine sulphate	150mg/kg	2	11.0 \pm 1.4	22.5 \pm 10.6
Control	0	6	12.3 \pm 2.3	10.7 \pm 1.6
Arecoline	1mg/kg	2	25.0 \pm 21.2	37.5 \pm 31.8
Arecoline	2mg/kg	2	11.0 \pm 1.4	14.0 \pm 8.5
Arecoline	4mg/kg	2	10.5 \pm 0.7	20.0 \pm 14.1

Table 4: Neurotransmitters and Metabolites mmoles/g in Mice Following, +35Gz For 30 sec.

Treatment	Number	NE	E	DA	DOPAC	HVA	5-HT	HIAA
Control 35 Gz/30 s	6	1.22 ± 0.22	0.69± 0.20	34.50± 2.20	3.30± 0.30	2.99± 0.50	14.12± 1.20	2.76± 0.30
	7	1.48 ± 0.33	0.62± 0.22	42.31± 9.30	3.30± 1.30	3.49± 0.60	16.52± 5.00	3.68± 0.90
Control 35 Gz/30 s	5	0.96 ± 0.17	0.77± 0.48	3.09 ± 0.50	0.24± 0.28	0.27± 0.11	5.34± 0.46	1.34± 0.12
	7	0.89 ± 0.32	0.34± 0.20*	2.16 ± 0.36	0.32± 0.27	0.35± 0.15	2.00±1.60*	1.95± 0.90
Control 35 Gz/30 s	6	1.62 ± 0.70	0.32± 0.13	1.31 ± 0.28	0.49± 0.19	0.32± 0.1	12.18±1.44	4.42± 0.50
	7	1.66 ± 0.40	1.17± 0.08*	1.37 ± 0.56	0.78± 0.30	0.55± 0.35	7.38±4.00*	5.14± 0.50*
Control 35 Gz/30 s	6	2.54 ± 0.26	0.67± 0.24	6.25 ± 1.60	0.88± 0.60	1.10± 0.30	17.11±2.00	5.14± 1.60
	7	2.34 ± 0.66	0.42± 0.20	5.69 ± 1.70	2.03± 0.90*	1.73± 0.70	7.60±2.40*	6.66± 1.60

(*) significantly different from control $p < 0.05$

The abbreviations are:

NE = Norepinephrine, E = Epinephrine, DA = Dopamine, HVA = Homovanillic acid, 5HT = 5 Hydroxytryptamine, DOPAC = 3,4 - dihydroxyphenyl acetic acid, 5-HIAA = 5- hydroxy indole acetic acid.

Table 5: The Effect in Mice of 0.2% Caffeine in Drinking Water For 8 Weeks on Neurotransmitters and Metabolites Following +35Gz For 30 Seconds

Treatment	Number	NE	E	DA	DOPAC	HVA	5-HT	HIAA
Cortex								
Control	4	1.45 ± 0.40	0.29 ± 0.33	13.27 ± 2.24	0.50 ± 0.60	3.65 ± 0.56	11.21±1.70	2.85± 0.44
35 Gz/30 s	2	1.45 ± 0.11	0.45 ± 0.06	13.59 ±3.20	2.17 ± 1.80	4.98 ± 1.60	9.39± 2.20	3.13± 1.00
Cerebellum								
Control	4	1.12 ± 0.90	0.60 ± 1.00	1.97 ± 0.70	0.07 ± 0.03	1.01 ± 0.50	5.81±0.50	2.36± 0.50
35 Gz/30 s	2	0.42 ± 0.20	0.91 ± 1.10	0.60 ± 0.80	0.47 ± 0.70	0.89 ± 0.45	1.18±0.11*	0.95± 0.37*
Brain stem								
Control	4	1.80 ± 0.70	0.49 ± 0.80	0.43 ± 0.30	0.58 ± 0.10	0.88 ± 0.30	4.24±2.11	3.31± 0.55
35 Gz/30 s	2	0.72 ± 0.80	0.29 ± 0.26	0.16 ± 0.22	0.12 ± 0.	2.97 ± 2.70	1.31±1.80	1.14± 1.60
MidBrain								
Control	4	1.88 ± 0.23	0.19 ± 0.25	5.76 ± 0.76		2.79 ± 0.36	9.25±2.40	4.42± 1.10
35 Gz/30 s	2	1.13 ± 1.40	0.59 ± 0.24	7.38 ± 3.20		3.35 ± 0.40	8.55±2.80	3.37± 0.08
Hippocamp								
Control	4	1.63 ± 0.32	0.72 ± 1.10	5.24 ± 1.20		2.06 ± 0.34	13.93±1.50	4.88± 2.00
35 Gz/30 s	2	1.70 ± 0.66	0.25 ± 0.10	6.99 ± 2.50		2.34 ± 0.16	14.43±4.00	3.51± 0.16

(*) significantly different from control $p < 0.05$

The abbreviations are:

NE = Norepinephrine, E = Epinephrine, DA = Dopamine, HVA = Homovanillic acid, 5HT = 5 Hydroxytryptamine, DOPAC = 3,4 - dihydroxyphenyl acetic acid, 5-HIAA = 5- hydroxy indole acetic acid

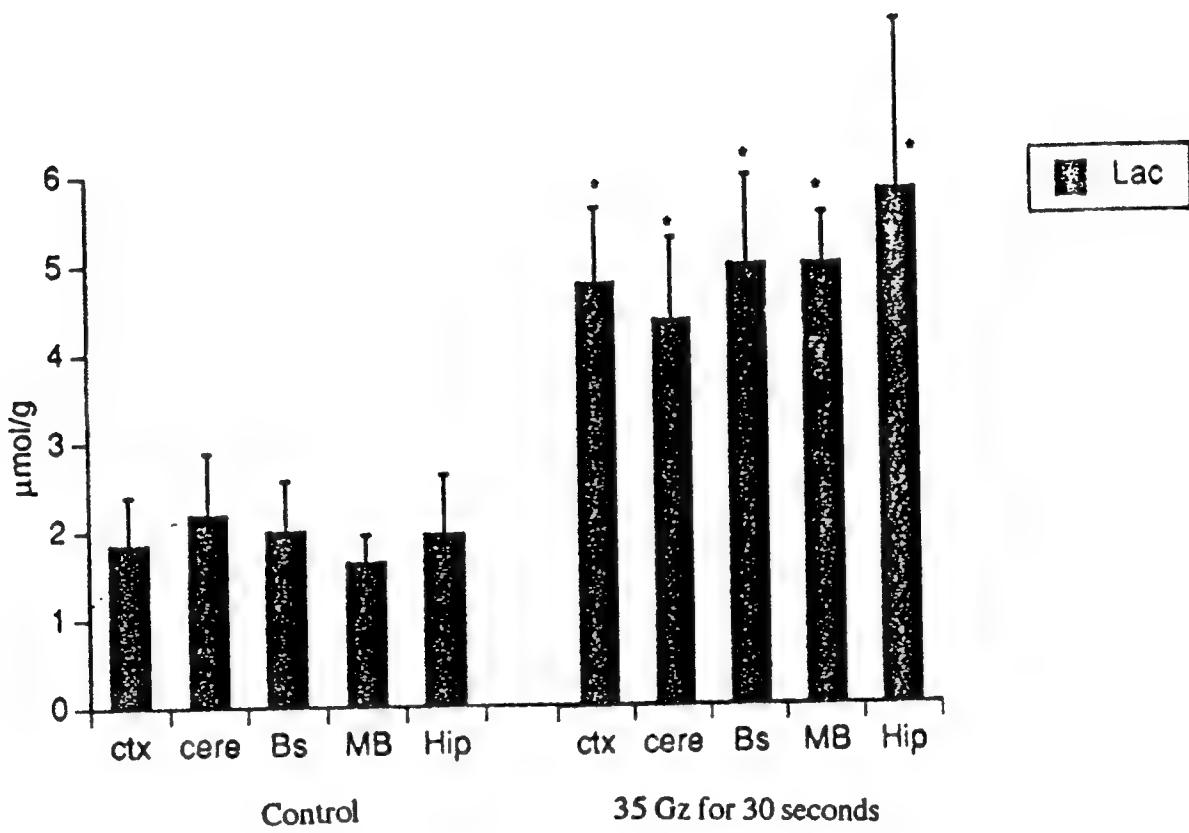


Fig 1: Lactate levels in $\mu\text{mol/g} \pm \text{sd}$ in mouse brain regions following 35 Gz for 30 seconds
 (*) Significantly different from control $p < 0.05$

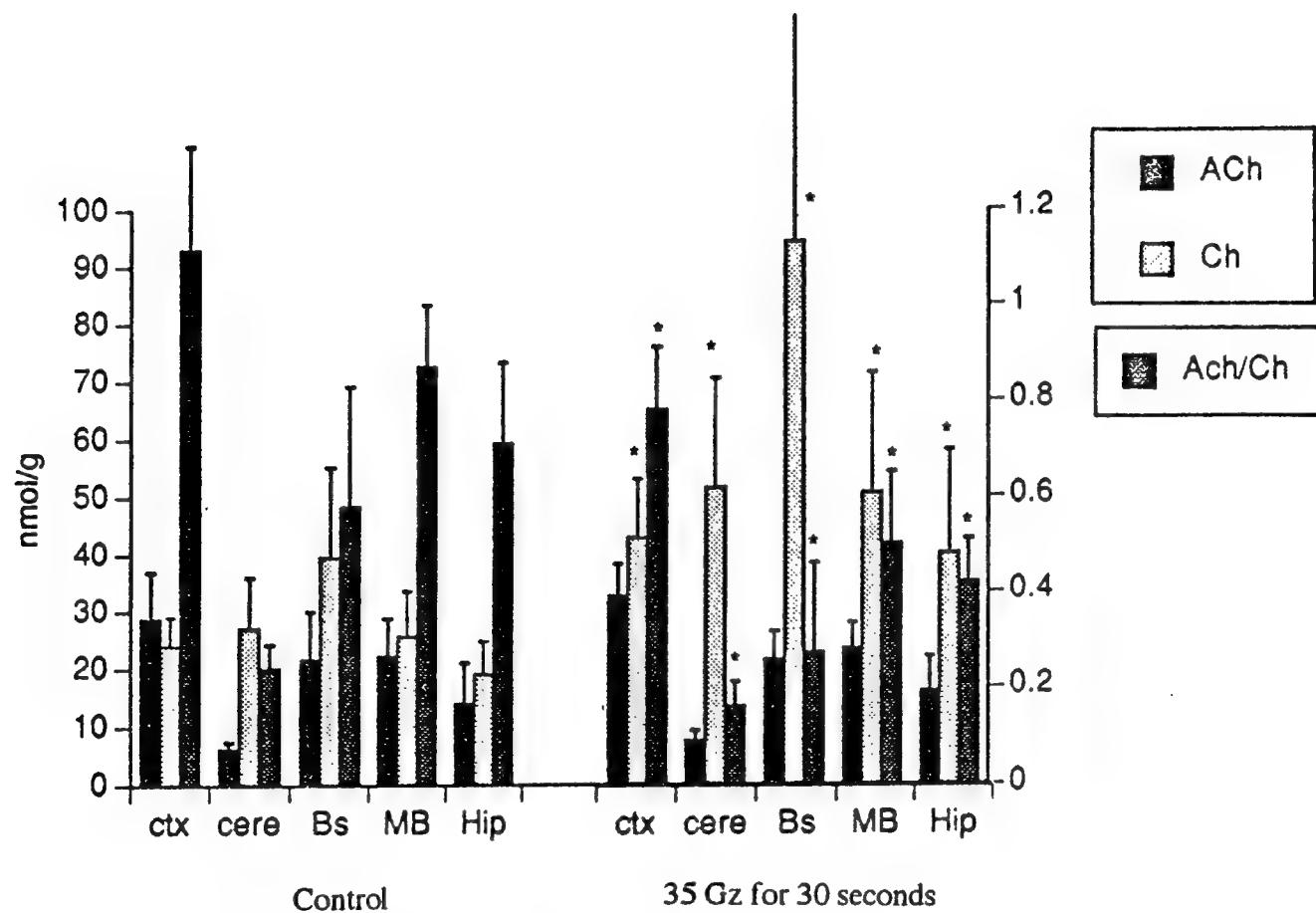


Fig 2: Acetylcholine and Choline levels in nmol/g mean \pm sd in mouse brain regions following 35 Gz for 30 seconds (*) significantly different from control $p < 0.05$

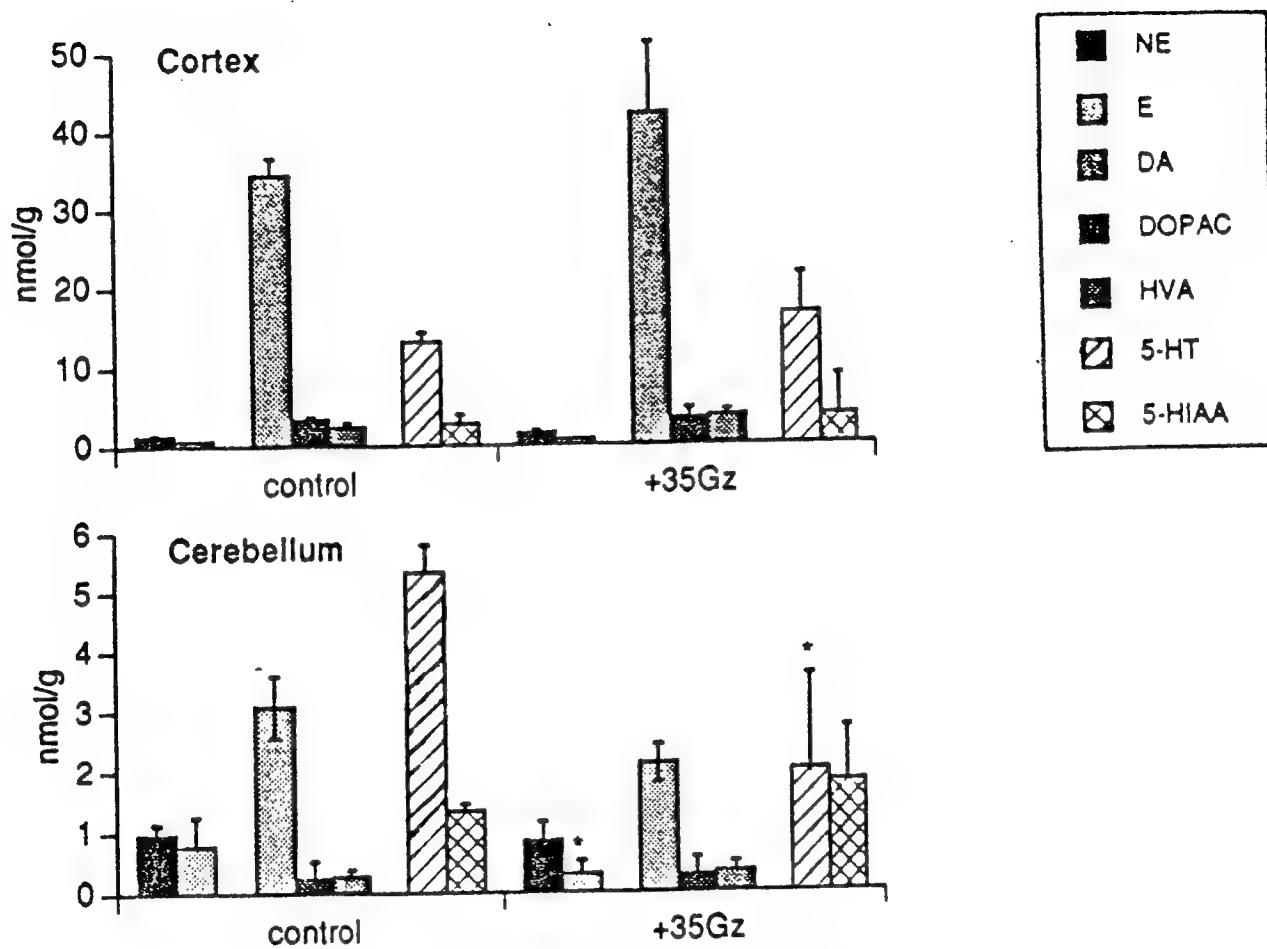


Fig 3a: Levels of neurotransmitters and metabolites in mouse brain: the effect of 30 seconds of 35 Gz.

The abbreviations are:

NE = norepinephrine, E = Epinephrine, DA = Dopamine
 HVA = Homovanillic acid, 5 HT = 5 - Hydroxytryptamine,
 DOPAC = 3, 4 dihydroxyphenyl acetic acid, 5-HIAA =
 5- hydroxy indole acetic acid.

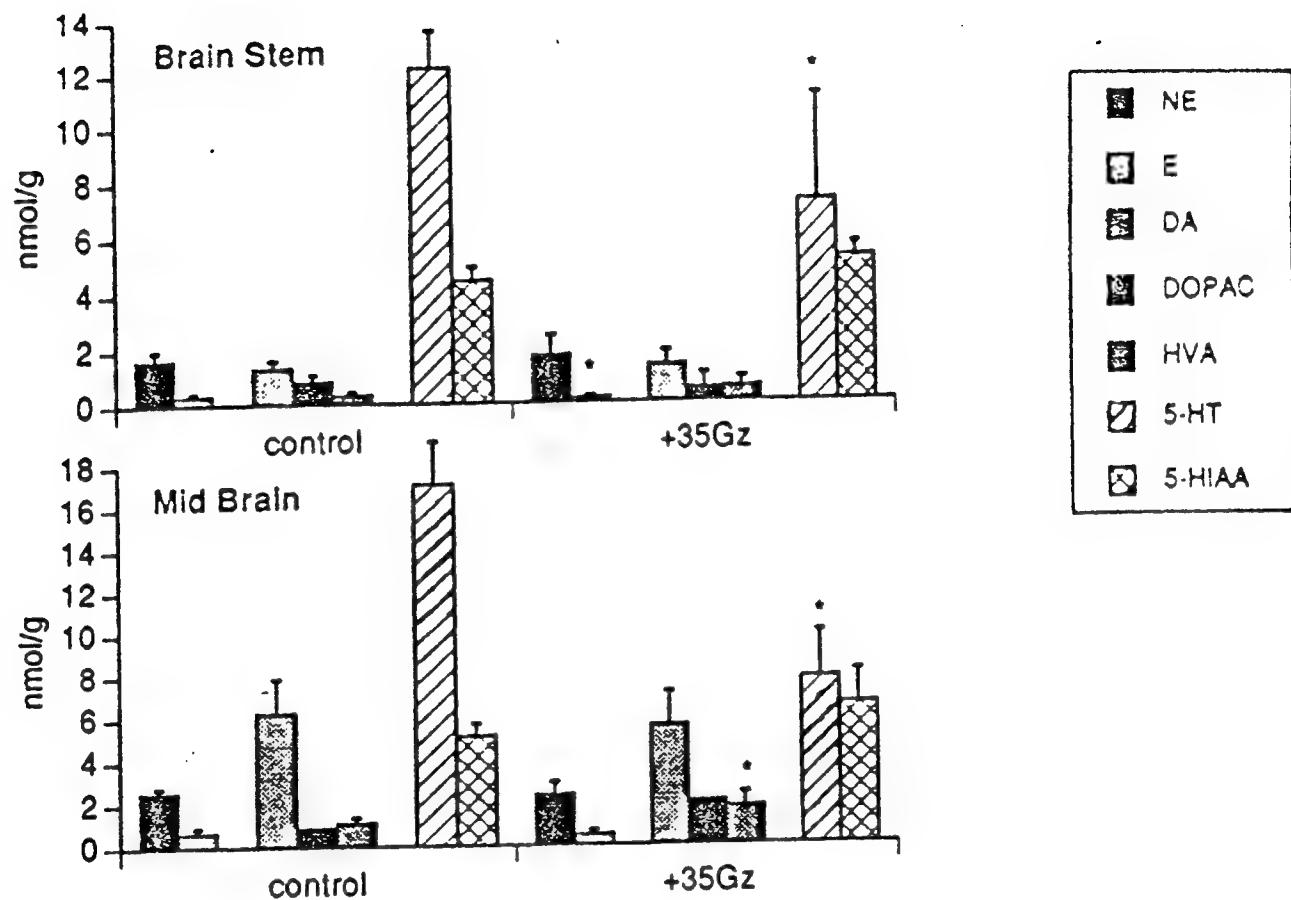


Fig 3b: Levels of neurotransmitters and metabolites in mouse brain: the effect of 30 seconds of 35 Gz.

The abbreviations are:

NE = norepinephrine, E = Epinephrine, DA = Dopamine
 HVA = Homovanillic acid, 5 HT = 5 - Hydroxytryptamine,
 DOPAC = 3, 4 - dihydroxyphenyl acetic acid, 5-HIAA =
 5- hydroxy indole acetic acid.

References

Astrup, J., Rehncrona, S., & Siesjo, B. K. (1980). The increase in extracellular potassium concentration in the ischemic brain in relation to preischemic functional activity and cerebral metabolic rate. *Brain Research*, 199, 161-174.

Benington, J. H., & Heller, H. C. (1995). Restoration of brain energy metabolism as the function of sleep. *Progress in Neurobiology*, 45, 347-360.

Forster, E. M., Cammarota, J. P., & Whinnery, J. E. (1994). G-LOC recovery with and without G-suit inflation. *Aviation, Space, and Environmental Medicine*, 65(March), 249-253.

Forster, E. M., & Whinnery, J. E. (1988). Recovery from Gz-induced loss of consciousness: psychophysiologic considerations. *Aviation, Space, and Environmental Medicine*, 59(June), 517-522.

Fredholm, B. B. (1995). Adenosine, adenosine receptors and the actions of caffeine. *Pharmacology and Toxicology*, 76, 93-101.

Freeman, G. B., Mykytyn, V., & Gibson, G. E. (1987). Differential alteration of dopamine, acetylcholine, and glutamine release during anoxia and/or 3,4-diaminopyridine treatment. *Neurochemical Research*, 12 (11), 1019-1027.

Gibson, G. E., & Blass, J.P. (1976). Impaired synthesis of acetylcholine in brain accompanying mild hypoxia and hypoglycemia. *Journal of Neurochemistry*, 27, 37-42.

Liptrot, J., Holdup, D., & Phillipson, O. (1993). 1,2,3,4,-tetrahydro-2-methyl-4,6,7-isoquinolinetriol depletes catecholamines in rat brain. *Journal of Neurochemistry*, 61, 2199-2206.

Peterson, C., Nicholls, D. G., & Gibson, G. E. (1985). Subsynaptosomal distribution of calcium during aging and 3,4-diaminopyridine treatment. *Neurobiology of Aging*, 6, 297-304.

Peterson, C., & Gibson, G. E. (1982). 3,4-Diaminopyridine alters acetylcholine metabolism and behavior during hypoxia. *The Journal of Pharmacology and Experimental Therapeutics*, 222(3), 576-582.

Phillis, J. W., & Wu, P. H. (1981). The role of adenosine and its nucleotides in central synaptic transmission. *Progress in Neurobiology*, 16, 187-239.

Prince, D. A., & Stevens, C. F. (1992). Adenosine decreases neurotransmitter release at central synapses. *Proc. Natl. Acad. Sci. USA*, 89(September), 8586-8590.

Rudolphi, K. A., Keil, M., Fastbom, J., & Fredholm, B. B. (1989). Ischaemic damage in gerbil hippocampus is reduced following upregulation of adenosine(A1) receptors by caffeine treatment. *Neuroscience Letters*, 103, 275-280.

Shahed, A. R., Barber, J. A., Galindo, S. J., & Werchan, P. M. (1995). Rat brain glucose and energy metabolites: effect of a +Gz (head to foot inertial load) exposure in a small animal centrifuge. *Journal of Cerebral Blood Flow and Metabolism*, accepted.

Shahed, A. R., Barber, J. A., & Werchan, P.M.(1994). Multiple +Gz exposures cause brain edema in rats. *Aviation, Space, and Environmental Medicine*, 65, 522-526.

Stavinoha, W. B. (1980). Acetylcholine in the brain during spontaneous and drug induced sleep. *Symposium on Cholinergic Mechanisms*, Florence, Italy.

Stavinoha, W. B. (1993). Use of microwaves for rapid fixation of tissues in vivo. *Scanning*, 15, 115-117.

Stavinoha, W. B., & Weintraub, S. T. (1974). Choline content of rat brain. *Science*, 183, 964-965.

Steriade, M. (1993). Cholinergic blockage of network- and intrinsically generated slow oscillations promotes waking and REM sleep activity pattern in thalamic and cortical neurons. *Progress in Brain Research*, 98, 345-355.

Steriade, M. (1993). Modulation of information processing in thalamocortical systems: chairman's introductory remarks. *Progress in Brain Research*, 98, 341-343.

Winn, R. H., Rubio, R., & Berne, R. M. (1979). Brain adenosine production in the rat during 60 seconds of ischemia. *Circulation Research*, 45 (4), 486-492.

Pentafluorobenzoylation and Mass Spectroscopic Analysis of Selected Amphetamines

John G. Stuart
Associate Professor
Department of Natural Sciences

Wiley College
Marshall, TX 75670

Final Report for:
Summer Faculty Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base
Washington, D.C.

and

Armstrong Laboratory

July 1995

**Pentafluorobenzoylation And Mass Spectroscopic
Analysis Of Selected Amphetamines**

**John G. Stuart
Associate Professor
Division of Natural Sciences
Wiley College**

Abstract

A selected series of amphetamine compounds (1-phenyl-2-aminopropanes) was pentafluorobenzoylated utilizing stoichiometric amounts of substituted amines (amphetamines or their hydrochlorides), pentafluorobenzoyl chloride, and triethylamine in dichloromethane at 65°C for 1 hour. The reaction mixtures, after washing with water and drying over anhydrous sodium sulphate, were evaporated under a stream of nitrogen using a Zymark Turbo LV Evaporator. The resulting solids or residues were reconstituted in ethyl acetate and the samples were subjected to mass spectroscopic analysis utilizing basically three methods: EI- Electron Ionization, PosCI- Chemical Ionization with positive ions and NICI- Negative Ion Chemical Ionization. Considering the information obtained, fragmentation mechanisms were proposed.

Pentafluorobenzoylation And Mass Spectroscopic Analysis Of Selected Amphetamines

John G. Stuart

The way mass spectra are deduced is the direct result of the chemistry of gaseous ions' activity in the ionization chamber of a mass spectrometer. After detection of the ions at the detector that measures the m/z ratio of each fragment, the spectral data are thus interpreted. The gaseous ions can be formed by a variety of methods and can be mass analyzed and detected instrumentally. As in many chemical reactions utilized for analysis, the purpose of the mass spectrometer is to convert the sample compound into quantifiable and measurable products that are related to the starting compound.

There are basically two types of ionization classes. Depending on the method used; gaseous positive ions are detected and the masses and relative abundances is displayed in the mass spectrum. Negative ions can give higher sensitivities for electronegative compounds; however, their fragmentations are not used much for structure determination.

Electron Ionization - The mass spectrum shows the mass of the molecule and the masses of fragments from it [1]. In this method, which is typically characterized as the "hard" ionization, at 70-eV, ionized molecules are produced which have high internal energies that they fragment before leaving the ion source. In this process the reagent producing the ionic products is a beam of energetic electrons.

"Soft" Ionization (Chemical Ionization) on the other hand is utilized such that the pressure in the ionization chamber is thus increased ~ 100 Pa (0.75 torr). This process typically utilizes as reagent gas methane or ammonia; and, it also assures that sample molecules will undergo hundreds of collisions before they escape the ionization chamber. Electron bombardment of methane, as an example, at such pressure in a closed ion source, produces reagent ions such as CH_5^+ (17) and C_2H_5^+ (29) [2]. These ions react with the sample through ion-molecule reactions. See Equations 1-5. Some stable products result and can be represented as undissociated sample molecules. With this methodology, different kinds of molecules can be ionized with specific positive and negative ions. Heteroatom-containing compounds such as the amphetamine derivatives described in this report can give abundant MH^+ ions in PosCl. Furthermore, the highly electronegative pentafluorobenzoyl moiety which contains an abundance of the very electronegative element fluorine on the benzene ring, is an ideal substrate to be utilized in the Negative Ion Chemical Ionization (NICI).

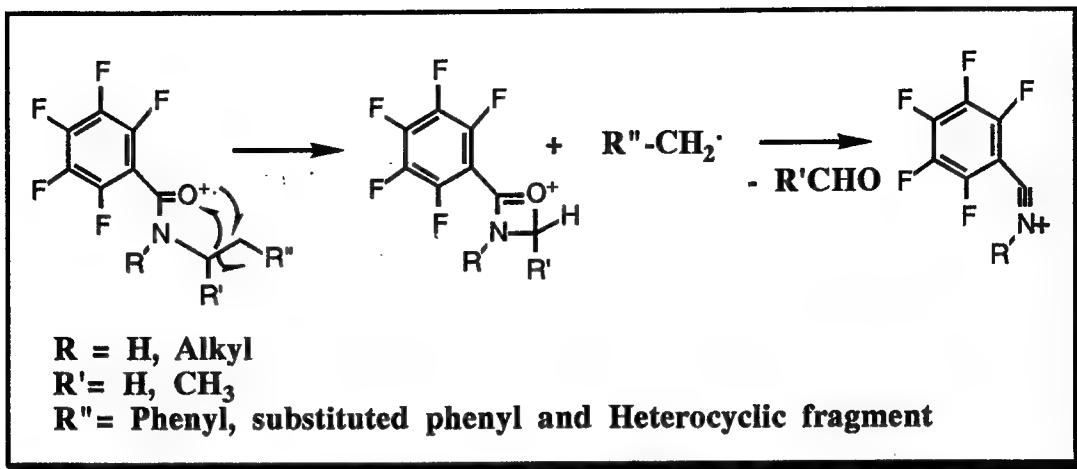


PosCI is very useful when molecular ions cannot be obtained in the EI methodology.

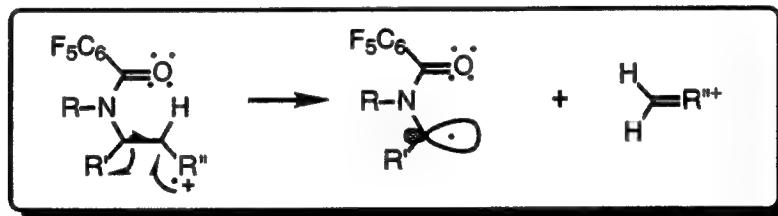
Amphetamine and its analogues are sympathomimetic phenethylamine derivatives with excellent central stimulant activity [3]. These compounds are useful in the treatment of obesity, narcolepsy and hypotension. For the most part these compounds are obtained in the dl-isomeric forms. These compounds are frequently abused for their stimulant effects and in most aspects the drugs are manufactured illegally and are therefore considered to be illicit if not properly prescribed by a physician. Furthermore, the abuse of stimulant drugs like amphetamine and its analogues has warranted interest in these compounds in forensic science and toxicology. This report describes the results of the mass spectroscopic analysis of a series of pentafluorobenzamides of amphetamine and its analogs. Three methods were utilized: EI- Electron Ionization, PosCI - Positive Ion Chemical Ionization and NICI- Negative Ion Chemical Ionization. We shall focus our discussion mostly on the first two methods.

The fragmentation of pentafluorobenzoyl derivatized of amphetamine and its analogues can be rationalized in three basic mechanistic schemes.

A first plausible mechanistic approach explains the formation of an azoxetonium ion, m/z (238, $\text{C}_9\text{H}_5\text{NOF}_5^{+}$) and (196, $\text{C}_7\text{HNF}_5^{+}$ which is not seen) fragments of amphetamine and the m/z (252, $\text{C}_{10}\text{H}_7\text{NOF}_5^{+}$) and (208, $\text{C}_8\text{H}_3\text{NF}_5^{+}$) of methamphetamine pentafluorobenzoylated derivatives.

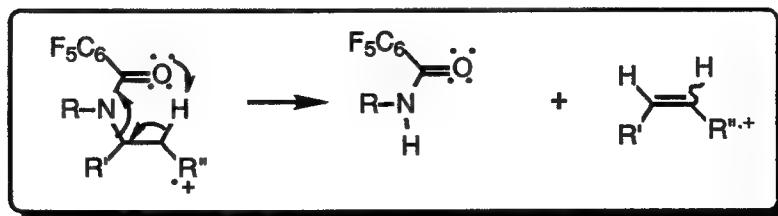


A second mechanistic approach is one which proceeds through the loss of an electron from the following radical-cationic compounds: benzyl, substituted phenyl, and a heterocyclic (indole derivatives, tryptamine and methyltryptamine) amphetamine derivatives. As a result, the following cationic fragments are potentially observed: m/z (91, $\mathbf{C}_7\mathbf{H}_7^+$, 125, $\mathbf{C}_7\mathbf{H}_6\mathbf{Cl}^+$; 151.2, $\mathbf{C}_9\mathbf{H}_{11}\mathbf{O}_2^+$; 108.9, $\mathbf{C}_7\mathbf{H}_6\mathbf{F}^+$; 121.2, $\mathbf{C}_8\mathbf{H}_9\mathbf{O}^+$; 129.9, $\mathbf{C}_9\mathbf{H}_8\mathbf{N}^+$) fragment.



$\mathbf{R} = \mathbf{H, Alkyl}$
 $\mathbf{R'} = \mathbf{H, CH}_3$
 $\mathbf{R''} = \mathbf{Phenyl, substituted phenyl and Heterocyclic fragment.}$

A third mechanistic possibility is the McLafferty rearrangement which explains the presence of m/z (152.2, $\mathbf{C}_9\mathbf{H}_9\mathbf{Cl}^+$; $\mathbf{C}_{11}\mathbf{H}_{14}\mathbf{O}_2^+$; 121.9, $\mathbf{C}_8\mathbf{H}_7\mathbf{F}^+$; 149.1, $\mathbf{C}_{10}\mathbf{H}_{13}\mathbf{O}^+$; 143, $\mathbf{C}_{10}\mathbf{H}_9\mathbf{N}^+$; 142.9, $\mathbf{C}_{10}\mathbf{H}_9\mathbf{N}^+$) fragment.



$\mathbf{R} = \mathbf{H, Alkyl}$

R' = H, CH₃

R'' = Phenyl, substituted phenyl and Heterocyclic fragment.

Experimental

General. All spectra were recorded on a Finnigan model 700 TSQ quadrupole mass spectrometer (Finnigan Corp., San Jose, CA) equipped with standard EI and CI source. Ionization of the reagent gas was accomplished using a 70-eV beam of electrons generated from a heated rhenium filament. The electron current was 200 uAmps for CI and 400 uAmps for EI. Reagent methane gas sample pressures was maintained at 8000 millitorr. Source temperatures were in the range of 100 - 150°C. Sample introduction was accomplished via a Varian Model 3000 series gas chromatograph interfaced to the ion source by means of a heated transfer line. Methane (99.999%) was purchased from Matheson, NJ.

Chemicals and Derivatives

Adamantadine.HCl, Pentafluorobenzoyl chloride, Phentermine.HCl, Fencamfamine.HCl, 4-bromo-2,5-dimethoxyphenethylamine.HCl, Mescaline, 2,4,5-trimethoxyphenethylamine.HCl, 4-methoxyamphetamine.HCl, 2,5-dimethoxyamphetamine.HCl, 3,4,5-trimethoxyamphetamine.HCl, 4-bromo-2,5-dimethoxyamphetamine.HCl, dibenzylamine, 1,2-diphenylethylamine, methyltryptamine, phenethylamine.HCl, tryptamine were purchased from Aldrich Chemical Co. Milwaukee, Wis. The deuterated compounds: *dl*-methamphetamine-D₈, *dl*-methamphetamine-D₁₁, *dl*-amphetamine-D₅, *dl*-amphetamine-D₈, *dl*-methylenedioxymethamphetamine-D₆, and *dl*-methamphetamine-D₆ were purchased from Radian Corp., Austin, TX.

Synthesis of Derivatives

Synthesis of 1-Adamantyl-pentafluorobenzamide as a typical example.

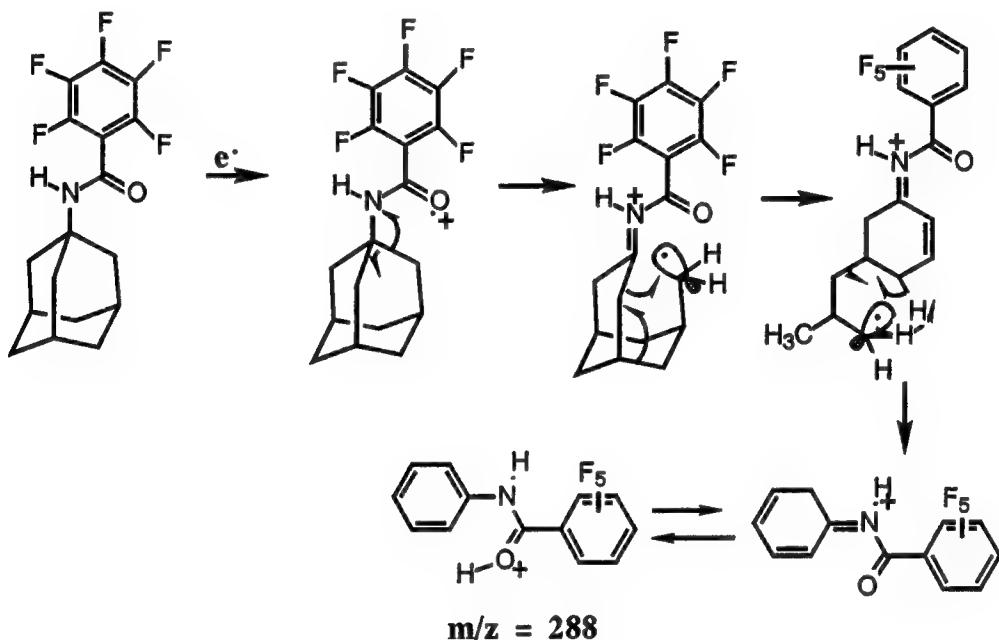
The hydrochloride of 1-amino adamantane (0.017g, 0.089 mmoles) was dissolved in 89.0 uL of 1N NaOH; from the basic solution, the free amine was extracted with 2.5 mL of methylene chloride. Triethylamine (12.5 uL, 0.089 mmoles) was added followed by pentafluorobenzoyl chloride (13 uL, 0.089 mmoles). The mixture was vortexed and allowed to reflux at 65°C for 1 hour. The solvent was evaporated and the resulting white solid was washed with water. The product was extracted with ethyl acetate 5.0 mL. After drying over sodium sulphate, the solvent was evaporated under a stream of nitrogen using a Zymark Turbo LV Evaporator. Further drying under vacuum resulted in a crystalline white solid.

Results and Discussion

The results of the methods used in the analysis of the derivatized amphetamines are consistent with the proposed mechanistic schemes that are aforementioned in the introductory discussion section. The results from some of the amphetamine analogs and the deuterated amphetamine analogs are outlined below.

Pentafluorobenzoyl chloride seems to be the most appropriate reagent for the amines described in this report for several reasons: the reaction is facile and the resulting pentafluorobenzamides are easily purified through recrystallization. A further significant reason for utilizing this reagent is that it contains a significant amount of fluorine on the benzene ring and the net result is that this reagent is very effective in electron capture detection.

Pentafluorobenzoyl-amantadine (C₁₇H₁₆NOF₅)

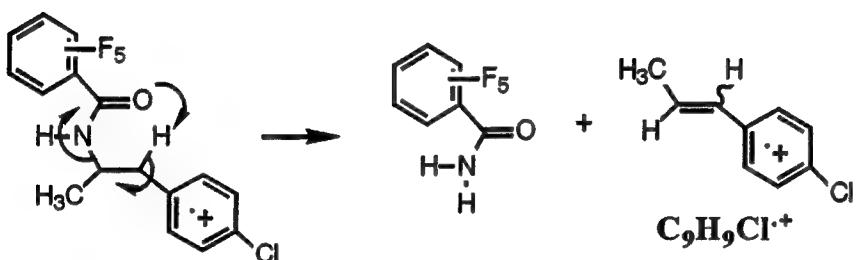
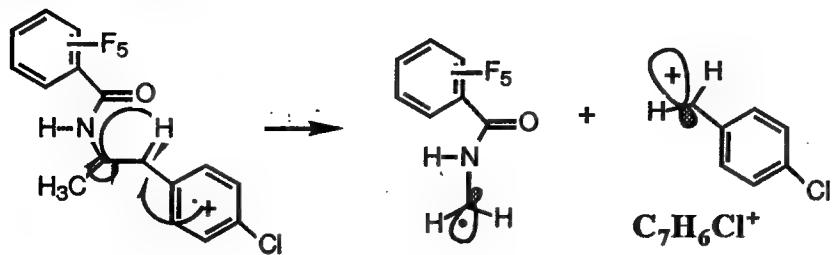


EI 344.9 (M^+ , $C_{17}H_{16}NOF_5$), 287.8 ($C_{13}H_6NOF_5^+$),
194.8 ($C_7F_5O^+$), 166.7 (C_6F_5), 92 ($C_6H_7N^+$)

PosCI 374.1 (M + 29, C₂H₅⁺), 346.1 (M+1), 345 (M⁺), 288 (C₁₃H₇NOF₅⁺), 194.5 (C₇F₅O⁺), 135 (Adamantyl-cation), 91 (C₇H₇⁺)

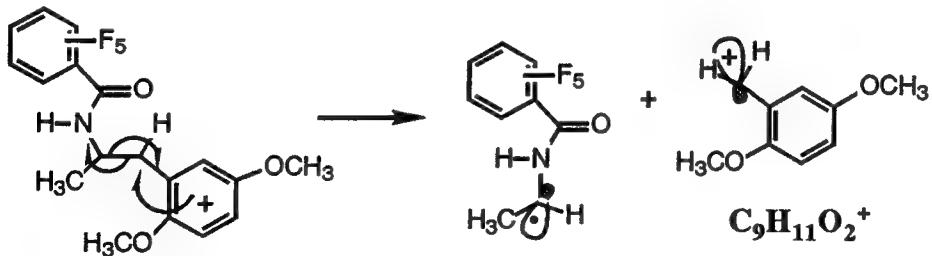
NICI 363.1 (M+18), 346 (M+1), 345 (M⁻), 325 (M-HF, 20)

Pentafluorobenzoyl-chloroamphetamine (C₁₆H₁₁NOClF₅)

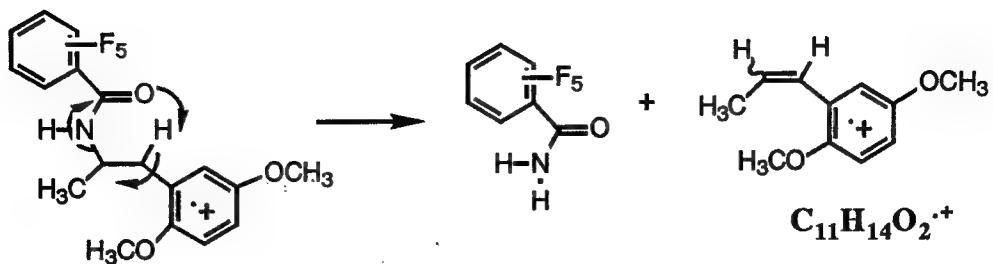


EI	238.2 ($C_9H_5NOF_5$), 195.1 ($C_7F_5O^+$), 167.1 ($C_9H_{10}NCl^+$), 152.2 ($C_9H_9Cl^+$), 125 ($C_7H_6Cl^+$)
PosCl	363.9 ($M+1$), 237.8 ($C_9H_5NOF_5^+$), 195.9 ($C_7F_5O^+$), 151.9 ($C_9H_9Cl^+$)
NICl	362.9 (M^-), 344.8 ($M-F, 19$), 322.8 ($M-2HF, 40$), 307 ($M-2HF-CH_3, 15$)

Pentafluorobenzoyl-2,5-dimethoxyamphetamine (C₁₈H₁₆NO₃F₅)



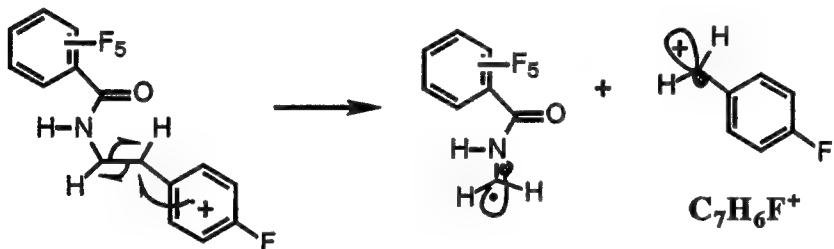
EI	389.3 (M^+), 238.2 ($C_9H_5NOF_5^+$), 195.1 ($C_7F_5O^+$), 178.2 ($C_{11}H_{14}O_2^+$), 151.2 ($C_9H_{11}O_2^+$)
----	--



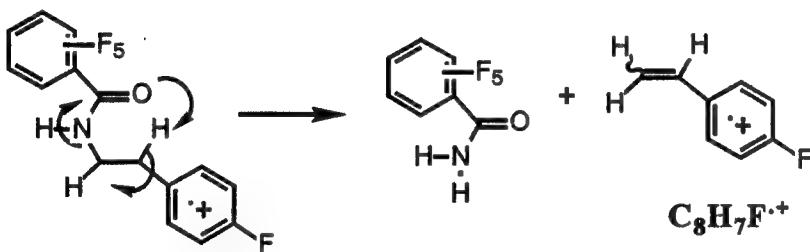
PosCI 418.2 (M+29, C_2H_5^+), 289.9 (M+1), 389 (M $^+$),
 237.9 ($\text{C}_9\text{H}_5\text{NOF}_5^+$), 194.8 ($\text{C}_7\text{F}_5\text{O}^+$),
 177.9 ($\text{C}_{11}\text{H}_{14}\text{O}_2^+$), 150.9 ($\text{C}_9\text{H}_{11}\text{O}_2^+$)

NICI 388.9 (M⁻), 370 (M-HF,20), 348.8 (M-3HF,40)

Pentafluorobenzoyl-p-fluorophenethylamine C₁₅H₁₀NOF₆



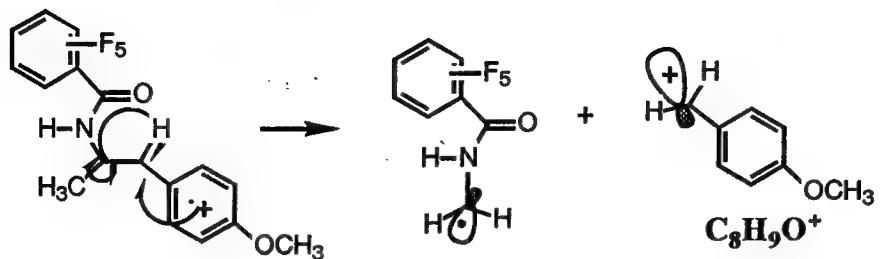
EI 194.8 ($\text{C}_7\text{F}_5\text{O}^+$), 121.9 ($\text{C}_8\text{H}_7\text{F}^+$), 108.9 ($\text{C}_7\text{H}_6\text{F}^+$)



PosCI **333.9 (M⁺), 239.9 (C₉H₇NOF₅⁺),
211.9 (C₇H₂NOF₅⁺), 194.9 (C₇F₅O⁺),
121.9 (C₈H₇F⁺), 109 (C₇H₆F⁺)**

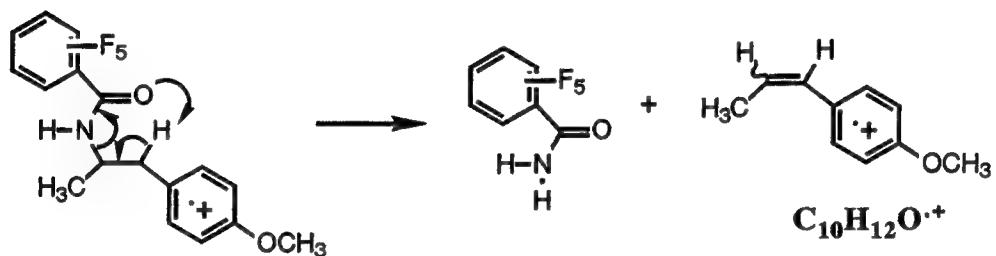
NICI 335.7 (M+2), 312.8 (M-HF,20), 292.9 (2HF,40),
147.7 (C₆F₄⁻)

Pentafluorobenzoyl-4-methoxyamphetamine ($C_{18}H_{16}NO_2F_5$)



EI

373.0 (M^+), 252.1 ($C_{10}H_7NOF_5^+$), 195.1 ($C_7F_5O^+$),
148.2 ($C_{10}H_{13}O^+$), 121.2 ($C_8H_9O^+$)



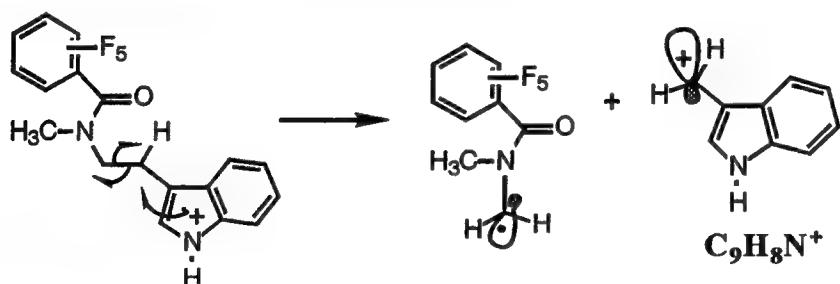
PosCI

374.1 ($M+1$), 252 ($C_{10}H_7NOF_5^+$), 195 ($C_7F_5O^+$),
149.1 ($C_{10}H_{13}O^+$), 121 ($C_8H_9O^+$)

NICl

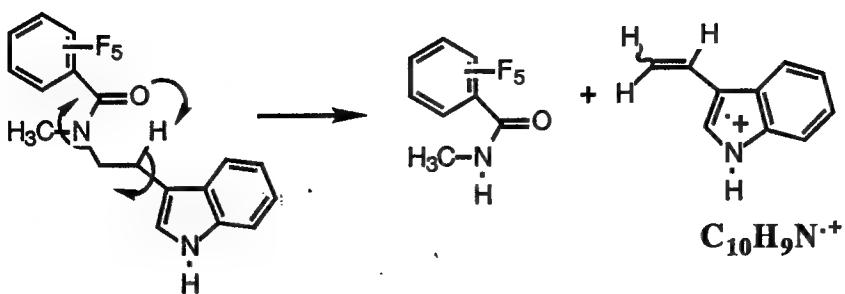
373 (M^-), 237 ($C_9H_4NOF_5^-$)

Pentafluorobenzoyl-methyltryptamine ($C_{18}H_{13}N_2OF_5$)



EI (ecurr,200)

367.9 (M^+), 194.8 ($C_7F_5O^+$), 142.9 ($C_{10}H_9N^{+\cdot}$),
129.9 ($C_9H_8N^+$)



EI (ecurr, 400) 367.9 (M^+), 194.8 ($C_7F_5O^+$), 143 ($C_{10}H_9N^{+ \cdot}$),
128.9 ($C_9H_8N^+$)

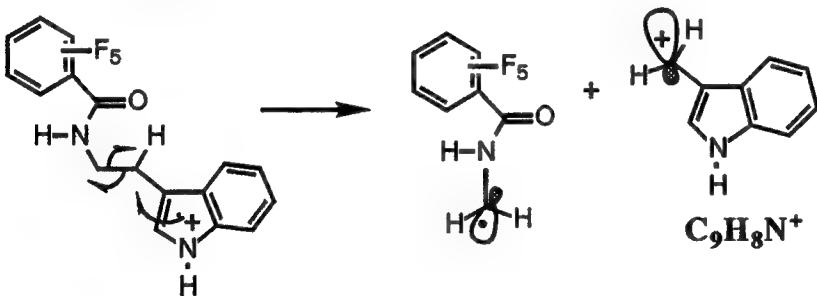
EI (ecurr, 800) 367.9 (M^+), 194.8 ($C_7F_5O^+$), 143 ($C_{10}H_9N^{+ \cdot}$),
129.8 ($C_9H_8N^+$)

PosCI 369.8 ($M+1$), 369 (M^+), 195 ($C_7F_5O^+$),
142.9 ($C_{10}H_9N^{+ \cdot}$), 129.9 ($C_9H_8N^+$)

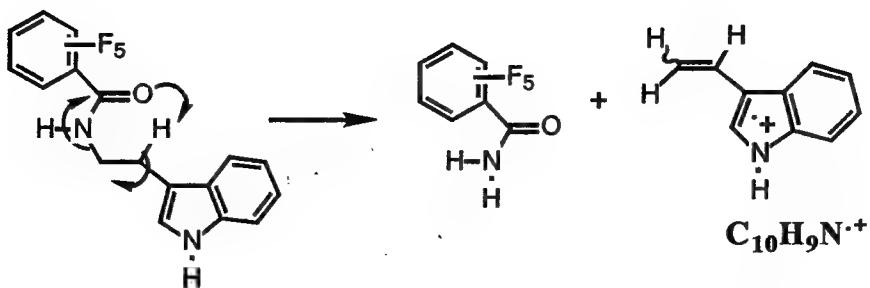
NegCI 347.8 ($M-HF$, 20), 327.8 ($M-2HF$, 40),
205.8 ($C_8HNF_5^-$)

NICI 348 ($M-F$, 19), 347.8 ($M-HF$, 20), 327.8 ($M-2HF$, 40),
205.8 ($C_8HNF_5^-$)

Pentafluorobenzoyl-tryptamine ($C_{17}H_{11}N_2OF_5$)



EI 353.9 (M^+), 195 ($C_7F_5O^+$), 142.9 ($C_{10}H_9N^{+ \cdot}$),
129 ($C_9H_8N^{+ \cdot}$)



PosCI	383.2 ($M + C_2H_5^+$, 29), 355.1 ($M + 1$), 195 ($C_7F_5O^+$), 172.1 ($C_{12}H_{14}N^+$), 143 ($C_{10}H_9N^+$), 130 ($C_9H_8N^+$)
NICl	335.1 ($M - F$, 19), 334 ($M - HF$, 20), 314 (2HF, 40), 294 ($M - 3HF$, 60), 205 ($C_8HNO_4^-$), 192 ($C_7H_2NOF_4^-$)
NegEI	333.6 ($M - HF$, 20), 313.5 ($M - 2HF$, 40), 204.8 ($C_8H_3NOF_4^-$), 192 ($C_7H_2NOF_4^-$)

The mass spectrum of the pentafluorobenzamide of 1-aminoadamantane shows the loss of 57, C_4H_9 , a butyl radical. Then a prominent ion peak with m/z of (288, $C_{13}H_7NOF_5^+$), is observed. Other significant ion fragments are the following: 195 ($C_6F_5O^+$), 344.9 (M^+), 166.8 ($C_6F_5^+$), 92 (Anilium ion). In the PosCI spectrum the following ions are observed: 135 (Adamantyl cation), 346.1 ($M + 1$), and 374.1 ($M + 29$, $C_2H_5^+$) and 194.5 ($C_6F_5O^+$). Surprisingly, in the NegEI spectrum, the molecular ion was indicated; however, what was common was the subsequent loss of F^- , HF and a molecular ion $C_{10}H_{14}N^-$, 148. The NICl spectrum also indicated a molecular ion 345 and 346 ($M + 1$) and the loss of (HF, 20) m/z 325.

Other pentafluoro benzamide compounds synthesized and analyzed in this study are the following: Pentafluorobenzoyl-Phentermine, $C_{17}H_{14}NOF_5$, MW = 343.30; Pentafluorobenzoyl-Fencamfamine, $C_{22}H_{20}NOF_5$, MW = 409.40; Pentafluorobenzoyl-4-bromo-2,5-dimethoxyphenethylamine, $C_{17}H_{13}NO_3BrF_5$, MW = 454.19; Pentafluorobenzoyl-Mescaline, $C_{18}H_{14}NO_4F_5$, MW = 403.31; Pentafluorobenzoyl-2,4,5-trimethoxyphenethylamine, $C_{18}H_{16}NO_4F_5$, MW = 405.32; Pentafluorobenzoyl-3,4,5-trimethoxyamphetamine, $C_{19}H_{18}NO_4F_5$, MW = 419.35; Pentafluorobenzoyl-4-bromo-2,5-dimethoxyamphetamine, $C_{18}H_{15}NO_3BrF_5$, MW = 468.22; Pentafluorobenzoyl-Phenethylamine, $C_{15}H_{10}NOF_5$, MW = 319.27; Pentafluorobenzoyl-Dibenzylamine, $C_{21}H_{14}NOF_5$, MW = 391.43 and Pentafluorobenzoyl-1,2-diphenylethylamine, $C_{21}H_{14}NOF_5$, MW = 391.34

The deuterated pentafluorobenzamides synthesized are the following:

Pentafluorobenzoyl-(*dl*)-Amphetamine-D₈ (C₁₇H₆D₈NOF₅, MW = 351.35);
Pentafluorobenzoyl-(*dl*)-Methamphetamine-D₁₁ (C₁₇H₂D₁₁NOF₅, MW = 353.35);
Pentafluorobenzoyl-(*dl*)-Amphetamine-D₅ (C₁₆H₇D₅NOF₅, MW = 334.30);
Pentafluorobenzoyl-(*dl*)-Amphetamine-D₈ (C₁₆H₄D₈NOF₅, MW = 337.32);
Pentafluorobenzoyl-(*dl*)-MDEA-D₆ (C₁₉H₁₀D₆NO₃F₅, MW = 407.37), and
Pentafluorobenzoyl-(*dl*)-1-Phenyl-2-methylaminopropane 1,1,2,3,3,3-D₆
(C₁₇H₈D₆NOF₅, MW = 349.34)

In examining the fragmentation pattern of the deuterated pentafluorobenzamides it is clear from fragments like:

EI: (334.20, C₁₆H₇D₅NOF₅⁺, (M⁺); 242, C₉HD₄NOF₅⁺; 195, C₇F₅O⁺; 167, C₆F₅⁺; 123.2, C₉H₅D₅⁺; 122, C₉H₆D₄⁺; and 92.1, C₇H₆D⁺

PosCI: 363.1 (M+29, C₂H₅⁺); 335.1 (MH⁺); 242, C₉HD₄NOF₅⁺; 195, C₇F₅O⁺; 124.1, C₉H₆D₅⁺, and 92.1, C₇H₆D⁺

NICI: 334.0 (M⁻); 313 (M-DF, 21); and 291.9 (M-2DF, 42) in pentafluorobenzoyl amphetamine-D₅.

The fragment ions for pentafluorobenzoyl amphetamine-D₈ are as follows:

EI: (337.5, C₁₆H₄D₈NOF₅⁺, (M⁺); 241.1, C₉HD₄NOF₅⁺; 195, C₇F₅O⁺; 167, C₆F₅⁺; 126.1, C₉H₂D₈⁺; 122, C₉H₆D₄⁺; and 96.1, C₇H₂D₅⁺

PosCI: 366.1 (M+29, C₂H₅⁺); 338.1 (MH⁺); 241.0, C₉HD₄NOF₅⁺; 126, C₉H₆D₈⁺; and 96, C₇H₂D₅⁺

NICI: 336.9 (M⁻); 318.1 (M-F, 19), and 317 (M-HF, 20), and 296 (M-HF (20) - DF (21) in pentafluorobenzoyl amphetamine-D₈.

The fragment ions for pentafluorobenzoyl methamphetamine-D₁₁ are as follows:

EI: (354.20, C₁₇H₃D₁₁NOF₅⁺, (M⁺); 258.1, C₁₀HD₆NOF₅⁺; 195, C₇F₅O⁺; 167, C₆F₅⁺; 126, C₉H₂D₈⁺; and 96.1, C₇H₂D₅⁺

PosCI: 383.1 (M+29, C₂H₅⁺); 355.1 (MH⁺); 258.1, C₁₀HD₆NOF₅⁺; 195, C₇F₅O⁺; 126, C₉H₂D₈⁺, and 96, C₇H₂D₅⁺

NICI: 354.0 (M⁻); 332.2 (M-DF, 21), and 317 (M-CH₃) in pentafluorobenzoyl amphetamine-D₁₁.

The fragment ions for pentafluorobenzoyl MDEA-D₆ are as follows:

EI: (407.2, C₁₉H₁₀D₆NO₃F₅⁺; M⁺; 272.1, C₁₁H₃D₆NOF₅⁺; 195, C₇F₅O⁺; 165.1, C₁₀H₇D₃O₂⁺; 135, C₈H₇O₂⁺

PosCI: 408.2 (MH⁺); 272.1, C₁₁H₃D₆NOF₅⁺; 194.9, C₇F₅O⁺; 165.1, C₁₀H₇D₃O₂⁺; 135, C₈H₇O₂⁺

NICI: 407.0 (M⁺) in pentafluorobenzoyl MDEA-D₆.

The m/z values of the deuterated pentafluorobenzamides are clearly indicative of the formation of the deuterated fragment ions such as the 3-N,N,N-trideuteromethyl-2-trideutero methyl- 4-pentafluorophenylazoxetonium, 258.1, C₁₀HD₆NOF₅⁺ cation, and the deuterated fragment, m/z 126, C₉H₂D₈⁺ from the McLafferty rearrangement. The values are consistent with the above proposed mechanisms.

Conclusion

In the present study, some pentafluorobenzamides of amphetamine analogs were examined through mass spectroscopy and their responses were examined using EI, CI and NICI. More emphasis was placed on the EI and CI methods because we were able to deduce more information about the structure of the compound from those two methods. The loss of F⁻, HF were consistently observed from the NICI spectra of the compounds studied when the NICI method was utilized. Aside from that, no appreciable bit of information was observed to allow us to know about the structural integrity of the compounds. The findings, spectral results from EI and CI, are consistent with the proposed mechanisms aforementioned earlier in this report.

Acknowledgements:

To God, my everlasting Father and Christ, my Savior and in the power of the Holy Spirit, I give thanks and praise. I am very greatful to the AFOSR and RDL for the opportunity to participate in the Summer Faculty Scientific Research. I am very greatful to Ms. Johnetta Thompson and Mr. Scott Liscoscos for their kindness and help during the program. I am also very greatful to Dr. Melvin S. Kaufman and his lovely wife, Col. Mehms, Maj Tacco and Alexander Hatzis and all the staff and support staff at AL/AOT for their wonderful support during the program. Finally, I am very greatful to my wife and three children for giving me the opportunity to leave home and to take part in this program.

References

- [1] Fred W. McLafferty and Frantisek Turecek, "Interpretation of Mass Spectra", 4th Ed.
- [2] Donald F. Hunt and Frank W. Crow, Analytical Chemistry, Vol. 50, No. 13 1978,
1781
- [3] O. Leroy Salerni, "Natural and Synthetic Organic Medicinal Compounds", 1976, C.V.
Mosby Co.

I have no inventions to report as a result of my work on the ASOFR 1995 Summer Research Program.

 08-16-95

Dr. John G. Stuart
95-0006

Regression to the Mean in Half-Life Studies

Ram C. Tripathi
Professor
Department of Mathematics and Statistics

Wiley College
Marshall, TX 75670

Final Report for:
Summer Faculty Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base
Washington, D.C.

and

Armstrong Laboratory

September 1995

REGRESSION TO THE MEAN IN HALF-LIFE STUDIES

Ram C. Tripathi

Professor

Department of Mathematics and Statistics
University of Texas at San Antonio

Abstract

Half-life studies of biomarkers for environmental toxins in humans are generally restricted to a few measurements per subject taken after the initial exposure. The initial dose is usually unknown because the exposure occurred before the substance was known to be toxic. In this setting, subjects are selected for inclusion in the study if their measured body burden is above a threshold (c) determined by the distribution of the biomarker in a control population. We assume a simple one-compartment first order decay model and a log-normal biomarker distribution, which together imply a repeated measures linear model relating the logarithm of the biomarker and time, with the slope being the negative of the decay rate (λ). Unless the data is properly conditioned, we show that ordinary weighted least-squares estimates of λ are biased due to regression toward the mean. These results are applied to show that published estimates of TCDD half-life in veterans of Operation Ranch Hand are biased. Based on these results an unbiased estimate is presented.

Key words: biomarker, half-life, regression toward the mean

REGRESSION TO THE MEAN IN HALF-LIFE STUDIES

Ram C. Tripathi

Introduction

Half-life studies of biomarkers for environmental toxins in humans are generally restricted to a few measurements per subject taken after the initial exposure. The initial dose is usually unknown because the exposure occurred before the substance was known to be toxic. We assume a one-compartment first order decay model with decay rate λ holds in subjects having body burdens above background levels defined by a threshold c . In this setting, subjects are selected for inclusion in the half-life study if their measured body burden is above c at some point in time. The threshold is defined as a high quantile, such as 0.99, of the biomarker distribution in a control population with no known exposure. We assume that the biomarker data is log-normally distributed, which together with the first order model implies a repeated measures linear model relating the logarithm of the biomarker and time, with the slope being $-\lambda$. Unless the data is properly conditioned prior to analysis, we show that the ordinary weighted least-squares estimate of λ is biased due to regression toward the mean.

These results are applied to a half-life study of 2,3,7,8 tetrachlorodibenzo-p-dioxin (dioxin) in veterans of Operation Ranch Hand (Michalek et al 1995). In section 2 we use the dioxin data to study the bias, variance and mean squared error of the estimates. An unbiased estimation procedure is described in section 3 and some remaining open issues are discussed in section 4.

Bias in the Estimation of the Decay Rate

We assume that k measurements per subject are available for each of n subjects, that the measurements were taken subsequent to the exposure and that the time between measurements (Δ) is fixed for all subjects, although the time between exposure and the first measurement for subject i (t_{i1}) varies between subjects. We assume that a single exposure produced an elevation of the body burden above background levels and that the first-order kinetics model

$$C_t = C_0 e^{-\lambda t} \quad (1)$$

holds in those subjects with body burdens above a threshold c . In (1), C_t is the concentration t years after exposure, C_0 is the (unknown) initial concentration, and λ is a constant but unknown decay rate. Based on (2.1), the population half-life is $t_{1/2} = \ln(2)/\lambda$. If we take the natural logarithm of (2.1), we obtain

$$\ln(C_t) = \ln(C_0) - \lambda t. \quad (2)$$

Thus, (2) can be regarded as a motivating equation for a linear model which accommodates multiple measurements per subject as well as covariates. Such a model, known as a fixed-effects repeated measures linear model is given by

$$y_{ij} = \mu + \tau_i + \beta t_{ij} + \varepsilon_{ij}, \quad (3)$$

for $i=1, 2, \dots, n$ and $j=1, 2, \dots, k$, where i indexes n subjects and j indexes k measurements per subject, y_{ij} represents the natural logarithm of the j th measurement on the i th subject,

$-\beta$ is the common decay rate (λ), $t_{ij}=t_{i1}+(j-1)\Delta$ is the time between exposure and the j th measurement, τ_i is the effect of the i th subject and ε_{ij} is the residual error term for y_{ij} .

Because the first order model (1) holds only in exposed subjects with body burdens above background levels, subjects are included in the half-life study only if $y_{ij} > \log(c)$, for all i and j . Given t_{ij} , $j=1, 2, \dots, k$, $\mathbf{y}_i=(y_{i1}, y_{i2}, \dots, y_{ik})'$ is k -variate normally distributed with covariance matrix Σ and mean $\mathbf{X}\beta$, where \mathbf{X} is the design matrix for (3) and $\beta'=(\mu, \beta, \tau_1, \tau_2, \dots, \tau_n)$. The weighted least-squares estimate of the vector of parameters β is

$$\mathbf{b}=(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y}, \quad (4)$$

where \mathbf{V} is a $kn \times kn$ block diagonal matrix with block Σ . When Σ is unknown, which is usually the case in practice, \mathbf{V} is replaced by its estimate \mathbf{S} , giving $\mathbf{b}=(\mathbf{X}'\mathbf{S}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{S}^{-1}\mathbf{y}$. Under the normal distribution assumption, \mathbf{b} is also the maximum likelihood estimate of β . In the special case that Σ is AR(1) and $k=3$, the weighted least-squares estimate of λ based on (4) is

$$\hat{\lambda} = \frac{1}{n} \sum_{i=1}^n \hat{\lambda}_{i13}, \quad (5)$$

the sample mean of the individual decay rates $\hat{\lambda}_{i13} = (y_{i1} - y_{i3})/2\Delta$.

Because the sample is truncated based on the selection criterion that all measurements are greater than $\log(c)$, $\hat{\lambda}$ will be biased due to regression to the mean, regardless of the sample size. For example, in the special case that $k=3$, and subject i is selected for the study if $y_{ij} > \log(c)$, $j=1,2,3$, $\hat{\lambda}$ will under-estimate the true decay rate λ . In this case (Tallis 1961),

$$E[\hat{\lambda} | y_{i1} > \log(c), y_{i2} > \log(c), y_{i3} > \log(c)] = \lambda - \text{BIAS},$$

where

$$\text{BIAS} = \frac{\sigma(1-\rho^2)}{2n\Delta} \sum_{i=1}^n \frac{1}{\alpha_i} [\phi(a_{i1})\Phi_2(A_{12i}, A_{13i}; \frac{\rho}{\sqrt{1+\rho^2}}) - \phi(a_{i3})\Phi_2(A_{31i}, A_{32i}; \frac{\rho}{\sqrt{1+\rho^2}})],$$

and

$$a_{i1} = \frac{\log(c) - E(y_{i1})}{\sigma}, \quad a_{i2} = \frac{\log(c) - E(y_{i2})}{\sigma}, \quad a_{i3} = \frac{\log(c) - E(y_{i3})}{\sigma},$$

$$E(y_{i1}) = \beta_0 + \tau_i + \beta_1(t_{i2} - \Delta), \quad E(y_{i2}) = \beta_0 + \tau_i + \beta_1 t_{i2}, \quad E(y_{i3}) = \beta_0 + \tau_i + \beta_1(t_{i2} + \Delta),$$

$$A_{12i} = \frac{a_{i2} - \rho a_{i1}}{\sqrt{1-\rho^2}}, \quad A_{13i} = \frac{a_{i3} - \rho^2 a_{i1}}{\sqrt{1-\rho^4}}, \quad A_{31i} = \frac{a_{i1} - \rho^2 a_{i3}}{\sqrt{1-\rho^4}}, \quad A_{32i} = \frac{a_{i2} - \rho a_{i3}}{\sqrt{1-\rho^2}}$$

$$\sigma^2 = \text{VAR}(y_{ij}), \quad j=1, 2, 3, \quad \rho = \text{corr}(y_{i1}, y_{i2}),$$

$$\alpha_i = P(x_{i1} > a_{i1}, x_{i2} > a_{i2}, x_{i3} > a_{i3}),$$

$x_{ij} = [y_{ij} - E(y_{ij})]/\sigma$, $i=1,2,\dots,n$, $j=1, 2, 3$, and ϕ is the standard normal density, Φ is the standard normal survival function. To express the effect of this bias, we compute the mean squared error (MSE), defined by $\text{MSE} = E(\hat{\lambda} - \lambda)^2 = \text{VAR}(\hat{\lambda}) + \text{BIAS}^2$.

In the same notation, the conditional variance of $\hat{\lambda}$ can be written as

$$\text{VAR}[\hat{\lambda} | y_{i1} > \log(c), y_{i2} > \log(c), y_{i3} > \log(c)] =$$

$$\frac{\sigma^2}{4n \Delta^2} \sum_{i=1}^n [\text{VAR}(x_{i1} | x_{i1} > a_{i1}, x_{i2} > a_{i2}, x_{i3} > a_{i3}) + \text{VAR}(x_{i3} | x_{i1} > a_{i1}, x_{i2} > a_{i2}, x_{i3} > a_{i3})]$$

$$-2\text{COV}(x_{i1}, x_{i3} | x_{i1} > a_{i1}, x_{i2} > a_{i2}, x_{i3} > a_{i3})],$$

where $\text{VAR}(x_{ij} | x_{i1} > a_{i1}, x_{i2} > a_{i2}, x_{i3} > a_{i3})$ and $\text{COV}(x_{i1}, x_{i3} | x_{i1} > a_{i1}, x_{i2} > a_{i2}, x_{i3} > a_{i3})$ are computed using the conditional moments in Tallis (1961) and these are given in the Appendix.

Example

The Air Force is conducting a 20-year prospective study of health and exposure to 2,3,7,8-tetrachlorodibenzo-p-dioxin (dioxin) in Air Force veterans of Operation Ranch Hand, the unit responsible for aerially spraying Agent Orange and other herbicides in Vietnam from 1962 to 1971 (Wolfe et al 1990). Comparison veterans who served in Southeast Asia during the same period but who were not involved with spraying herbicides serve as referents. Physical examinations were administered in 1982, 1985, 1987 and 1992. Since 1987, exposure has been indexed by a measurement of dioxin in serum. In 1987, all willing Ranch Hand veterans and Comparisons were asked to contribute blood for a dioxin assay; 889 Ranch Hands and 1,132 were assayed and

received a quantifiable result (Roegner et al 1991). As part of a study of dioxin half-life, all 343 Ranch Hand veterans with dioxin levels above 10 parts per trillion (ppt) in 1987 and who had stored serum from the 1982 examination were selected. Hence, $c=10$ ppt in this study. The 10 ppt cut point is the 98th percentile of the Comparison dioxin distribution and is the value we regard as the upper threshold for background exposure. Of the 343 subjects, 278 attended the 1992 examination and received a quantifiable dioxin result (Michalek et al 1995). These data, after log transformation, are summarized in Table 1. In summarizing the total population in 1987, we excluded 9 Ranch Hands and 76 controls with nondetectable dioxin levels to preserve the normality of the log transformed distributions.

Table 1
 Summary of Log Transformed Dioxin Distributions
 in the Air Force Health Study

Ranch Hand

Year	Half Life Study		Total		Comparison	
	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)
1982	343	3.8 (0.7)				
1987	343	3.6 (0.8)	880	2.68 (1.07)	1,056	1.46 (0.46)
1992	278	3.1 (0.8)				

In this example, 240 subjects had all three dioxin measurements above the threshold of $c=10$ ppt, giving $\hat{\lambda} = 0.077634$, $BIAS=0.002496$, $VAR(\hat{\lambda})=3.95445\times 10^{-5}$, and $MSE=4.57756\times 10^{-5}$. The correlation matrix of y_{i1}, y_{i2}, y_{i3} , given that $y_{ij} > \log(10)$ is

$$\begin{bmatrix} 1 & 0.86391 & 0.78594 \\ 0.86391 & 1 & 0.83871 \\ 0.78594 & 0.83871 & 1 \end{bmatrix}.$$

To compute the correlation of the untruncated data, we used the algorithm of Mee and Chua (1991), giving $\hat{\rho} = 0.88497$.

Unbiased Estimation of the Decay Rate

Inspection of (5) shows that the bias can be made equal to zero by truncating the data such that $a_{1i}=a_{i2}=a_{i3}$, or, equivalently, such that the observations lie above a straight line with slope $-\lambda$. This requires an iterative procedure, because λ is unknown. The first step is to truncate the data at $\log(c)$ at each time point and estimate λ using weighted least-squares with AR(1) correlation structure. The second step is to truncate the first measurement at $\log(c) - 2\hat{\lambda}\Delta$, the second at $\log(c) - \hat{\lambda}\Delta$ and the third at $\log(c)$. The third step is to repeat the process using the updated value of $\hat{\lambda}$. The process is repeated until the estimated value of λ is equal to the value used for truncation. The final estimate of λ is then unbiased.

When applied to the Air Force data, the procedure converged in two steps. At the first step, $\hat{\lambda} = 0.077637$, based on 240 subjects with all three measurements greater than $\log(10)$. In the second step, $\hat{\lambda} = 0.081188$, with $\text{VAR}(\hat{\lambda}) = 4.43057 \times 10^{-5}$, based on 213 subjects with $\Delta=5$ years and, in original units, with the first dioxin level above 22.5 ppt, the second above 15.0 ppt and the third above 10 ppt.

Extensions

There are many possible extensions to consider when attempting to generalize the material in the previous two sections. We have considered larger values of k under the AR(1) model and under the more general Toeplitz covariance model. These results are summarized below.

k=4, AR(1)

In this case, the weighted least-squares estimate of the decay rate can be written as

$$\hat{\lambda} = \sum_{i=1}^n (w_{14} \hat{\lambda}_{14i} + w_{23} \hat{\lambda}_{23i}),$$

where $w_{14} = \frac{3(3-\rho)}{n(10-5\rho+\rho^2)}$, $w_{23}=1-w_{14}$, $\hat{\lambda}_{14i} = (y_{i1} - y_{i4})/3\Delta$ and $\hat{\lambda}_{23i} = (y_{i2} - y_{i3})/\Delta$.

k=3, Toeplitz

In this case, $\hat{\lambda}$ is the same as that given by equation (5).

k=4, Toeplitz

In this case, $\hat{\lambda}$ has the same form as in the case $k=4$, AR(1), with weights

$$w_{14} = \frac{3(\theta_3 - 4\theta_2 + 3\theta_1)}{(15\theta_2 + \theta_4 - 6\theta_3 - 10\theta_1)} \text{ and } w_{23}=1-w_{14}, \text{ where}$$

$$\Sigma = \begin{bmatrix} \theta_1 & \theta_2 & \theta_3 & \theta_4 \\ \theta_2 & \theta_1 & \theta_2 & \theta_3 \\ \theta_3 & \theta_2 & \theta_1 & \theta_2 \\ \theta_4 & \theta_3 & \theta_2 & \theta_1 \end{bmatrix}.$$

In all of these cases, $\hat{\lambda}$ is biased if the data is truncated to lie above the same constant every point in time and is unbiased if the data is truncated to lie above a straight line with slope $-\lambda$.

We have been unable to prove this result for general values of k under either the AR(1) or Toeplitz covariance assumption, however, we conjecture that these results hold in general under these two covariance structures. However, we have proved that these results do not hold when Σ is equivariant ($\sigma_j^2 = \sigma_k^2, j \neq k, \sigma_{jk}$ arbitrary).

References

Tallis, G.M. (1961). The moment-generating function of the truncated multinormal distribution. Journal of the Royal Statistical Society B, 23:223-229.

Mee, T.W. and Chua, T.C. (1991). Regression toward the mean and the paired sample t test. The American Statistician 45:39-42.

Michalek, J.E., Pirkle, J.L., Caudill, S.P., Tripathi, R.C., Patterson Jr, D.G. and Needham, L.L. (1995). Pharmacokinetics of TCDD in veterans of Operation Ranch Hand: 10-year follow-up. Journal of Toxicology and Environmental Health (in press).

Wolfe, W.H., Michalek, J.E., Miner, J.C., Rahe, A.J., Silva, J., Thomas, W.F., Grubbs, W.D., Lustik, M.B., Garrison, T.G., Roegner, R.H., and Williams, D.E. (1990). Health status of Air Force veterans occupationally exposed to herbicides in Vietnam. JAMA 264:1824-1831.

Roegner, R.H., Grubbs, W.D. Lustik, M.B., Brockman, A.S., Henderson, S.C., Williams, D.E., Wolfe, W.H., Michalek, J.E. and Miner, J.C. (1991). The Air Force Health Study. An epidemiologic investigation of health effects in Air Force personnel following exposure to herbicides. Serum dioxin analysis of 1987 examination results. National Technical Information Service (AD A 237 516-524) Springfield, Virginia.

Appendix

$$\alpha_i E(x_{i1}) = \phi(a_{i1})\Phi(A_{i12}, A_{i13}; r) + \rho\phi(a_{i2})\Phi(A_{i21}, A_{i23}; 0) + \rho^2\phi(a_{i3})\Phi(A_{i31}, A_{i32}; r)$$

$$\alpha_i E(x_{i2}) = \rho\phi(a_{i1})\Phi(A_{i12}, A_{i13}; r) + \phi(a_{i2})\Phi(A_{i21}, A_{i23}; 0) + \rho\phi(a_{i3})\Phi(A_{i31}, A_{i32}; r)$$

$$\alpha_i E(x_{i3}) = \rho^2\phi(a_{i1})\Phi(A_{i12}, A_{i13}; r) + \rho\phi(a_{i2})\Phi(A_{i21}, A_{i23}; 0) + \phi(a_{i3})\Phi(A_{i31}, A_{i32}; r)$$

$$\alpha_i E(x_{i1}^2) = \alpha_i + a_{i1}\phi(a_{i1})\Phi(A_{i12}, A_{i13}; r) + \rho^2 a_{i2}\phi(a_{i2})\Phi(A_{i21}, A_{i23}; 0) + \rho^4 a_{i3}\phi(a_{i3})\Phi(A_{i31}, A_{i32}; r)$$

$$+ \rho\phi(a_{i1}, a_{i2})\Phi(A_{i13}^2)(1 - \rho^2) + \rho^2\phi(a_{i1}, a_{i3})\Phi(A_{i12}^3)(1 - \rho^2) + \rho^3\phi(a_{i2}, a_{i3})\Phi(A_{i21}^3)(1 - \rho^2)$$

$$\alpha_i E(x_{i3}^2) = \alpha_i + \rho^4 a_{i1}\phi(a_{i1})\Phi(A_{i12}, A_{i13}; r) + \rho^2 a_{i2}\phi(a_{i2})\Phi(A_{i21}, A_{i23}; 0) + \rho^4 a_{i3}\phi(a_{i3})\Phi(A_{i31}, A_{i32}; r)$$

$$+ \rho^2\phi(a_{i1}, a_{i2})\Phi(A_{i23}^1) + \rho^2(1 - \rho^2)\phi(a_{i2}, a_{i3})\Phi(A_{i21}^3) +$$

$$+ \rho\phi(a_{i1}, a_{i3})\Phi(A_{i32}^1)(1 - \rho^2) + \rho^3\phi(a_{i1}, a_{i3})\Phi(A_{i12}^3)(1 - \rho^2)$$

$$\alpha_i E(x_{i1} x_{i3}) = \rho^2 \alpha_i + \rho^2 a_{i1}\phi(a_{i1})\Phi(A_{i12}, A_{i13}; r) + \rho^2 a_{i2}\phi(a_{i2})\Phi(A_{i21}, A_{i23}; 0)$$

$$+ \rho^3 a_{i3}\phi(a_{i3})\Phi(A_{i31}, A_{i32}; r) + \rho(1 - \rho^2)\phi(a_{i1}, a_{i2})\Phi(A_{i23}^1) + (1 - \rho^4)\phi(a_{i1}, a_{i3})\Phi(A_{i32}^1)$$

$$+ \rho(1 - \rho^2)\phi(a_{i2}, a_{i3})\Phi(A_{i31}^2),$$

where $r = \rho / \sqrt{1 + \rho^2}$, $\phi(x)$ is the standard normal density, $\Phi(x)$ is the standard normal survival function, $\phi(x,y) = \phi(x,y; r)$ is the bivariate standard normal density with correlation r , $\Phi(x,y; r)$ is the bivariate upper tail probability of the bivariate standard normal distribution with correlation r , $A_{i23}^1 = A_{i13}^2$, $A_{i12}^3 = A_{i32}^1$, $A_{i21}^3 = A_{i31}^2$ and

$$A_{i13}^2 = \frac{a_{i3} - \rho a_{i2}}{\sqrt{1 - \rho^2}}, \quad A_{i12}^3 = \left(a_{i2} - \frac{\rho a_{i3}}{1 + \rho^2} - \frac{\rho a_{i1}}{1 + \rho^2} \right) \sqrt{\frac{1 + \rho^2}{1 - \rho^2}}, \quad A_{i31}^2 = \frac{a_{i1} - \rho a_{i2}}{\sqrt{1 - \rho^2}}.$$

Evaluation of a Revised Taxonomy in Research on Cross-Job Transferability of Skills

Stephen A. Truhon
Associate Professor
Department of Social Sciences

Winston-Salem State University
Winston-Salem, NC 27110

Final Report for:
Summer Faculty Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base
Washington, D.C.

and

Armstrong Laboratory

August 1995

EVALUATION OF A REVISED TAXONOMY IN RESEARCH ON
CROSS-JOB TRANSFERABILITY OF SKILLS

Stephen A. Truhon
Associate Professor
Department of Social Sciences
Winston-Salem State University

ABSTRACT

An examination was made of a revised 45-category taxonomy of tasks performed by Air Force personnel. Retrainees, supervisors and co-workers of these retrainees, as well non-retrainees representing Air Force specialties from which the retrainees were transferred participated in the study. They completed a questionnaire asking three or five questions about each of these task categories as well as other questions. Results showed the taxonomy has face validity, is reliable, is distinctive, and is useful in conducting research on cross-job transferability of skills. However, the reliability of several of the items as well as the low response rate suggest that this taxonomy be shortened in future studies.

EVALUATION OF A REVISED TAXONOMY IN RESEARCH ON CROSS-JOB TRANSFERABILITY OF SKILLS

Stephen A. Truhon
Associate Professor
Department of Social Sciences
Winston-Salem State University
601 Martin Luther King Jr. Drive
Winston-Salem, NC 27110

INTRODUCTION

The issue of cross-job transferability of skills has not been examined in great detail, because, until recently, it was assumed that most job transfers would be vertical, reflecting the promotions that employees receive. But in recent years increased interest has occurred in lateral job transfers because of rapid technological changes, mergers, and acquisitions. This issue has been especially important in the Armed Forces as the number of personnel decline and jobs need to be combined.

Theoretically, it has been assumed that transferability of skills from one job to another is related to the time necessary for retraining (Fine, 1957a; 1957b). Research has confirmed this (Gordon & Fitzgibbons, 1982; Lance, Kavanagh, & Gould, 1993; Sparrow, 1989). Typically this research has used decomposed estimates of retraining time (i.e., new and old jobs are broken down into a series of components and compared), but global estimates have been used effectively as well (Lance, Mayfield, Gould, & Linsky, 1991).

If jobs are to be analyzed by their components, the question of what taxonomy of tasks is to be used. Taxonomies have been used in many work settings, some fairly general while others specific to a work setting. The Air Force has attempted several such taxonomies. Perhaps the oldest is the MAGE system of classifying Air Force specialties into Mechanical, Administrative, General, and Electronic. While useful, the MAGE system has been viewed as "too broad to adequately describe the components or work characteristics of all jobs or specialties" (Bell & Thomasson, 1984, p.1). As a result several other taxonomies have been used: 1) a 26-category taxonomy devised by the United States Air Force Occupational Measurement Center (Bell & Thomasson, 1984); 2) a 33-category taxonomy derived from the General Work Inventory (Ballentine & Cunningham, 1981); and 3) five-component taxonomy derived from factor analysis of the Electronics Principles Inventory (Ruck, 1986).

Lance, Gould and their associates (Lance et al. 1991; Lance, Kavanagh, & Gould, 1993; Lance, Mayfield, & Gould, 1993; 1995) have used the above three taxonomies to develop a 26-category taxonomy that has proved useful in describing Air Force Specialties (AFS's) and supporting research on cross-job transferability of skills.

Cross-job transferability of skills may be viewed as a way of measuring ease of movement between jobs. Research (Lance, Kavanagh, & Gould, 1993; Truhon, 1993) suggests that two factors—job difficulty and job similarity— are related to ease of movement between jobs. More specifically, Lance, Kavanagh, and Gould (1993) found that greater ease of movement between jobs occurred when: 1) old jobs had higher levels of job difficulty; 2) new jobs had lower levels of job difficulty; and 3) the two jobs required similar aptitudes. Likewise, Truhon (1993) found that four variables, two dealing with job difficulty and two with job similarity, accounted for 30 to 32 percent of the variance in ease of movement.

METHOD

Sample

Four samples were examined in this study. A group of 3,000 Air Force personnel were selected who met the following criteria: 1) they had been retrained between May 1992 and April 1994; 2) they had been retrained into an AFS in which at least 100 airmen had been retrained during the same period; and 3) they were between an E-3 and E-6 level. Co-workers and supervisors to these retrainees made up the second and third samples. The fourth group consisted of 1,500 Air Force personnel who had not been retrained within the previous four years and were either from the same AFS's as the retrainees started from or retrained to. A list of the AFS's included in this study are presented in Table 1.

The total number of retrainees who returned questionnaires was 934. Among the co-workers 650 responded, but 184 could not be matched to retrainees, leaving 466. Of the supervisors, 968 responded, but 325 could not be matched to retrainees, leaving 643. The total number of non-retrainees was 530. There were further drops in the numbers of subjects (see Outliers).

Task Taxonomy

Lance, Kavanagh and Gould (1995) had previously devised a 26-category taxonomy for work on cross-job transferability of skills. They have revised this taxonomy to include 45 categories, but five categories (Medical-Technical, Animal Care, Nondestructive Aircraft Structure Inspection, Airborne Telemetry Monitoring and Maintenance, and Security Perimeter Maintenance) were dropped because fewer than 10 retrainees, in the current study, indicated that these activities were part of their job (Stennett, personal communication). The remaining 40 categories are presented in Table 2.

Table 1
Specialties Included in the Present Study

AFSC	U.S. Air Force Job Title
113X0C	Flight Engineer
114X0	Aircraft Loadmaster
201X0	Intelligence Operations Specialist
201X1	Target Intelligence Specialist
205X0	Electronics Intelligence Operations Specialist
241X0	Safety Specialist
242X0	Disaster Preparedness Specialist
251X0	Weather Specialist
271X2	Operations Resources Management Specialist
272X0	Air Traffic Control Operator
274X0	Command and Control Specialist
277X0	Space Systems Operations Specialist
304X2	Meteorological and Navigation Systems Specialist
304X4	Ground Radio Communications Specialist
304X5	Television Systems Specialist
304X6	Satellite Communications Systems Equipment Specialist
305X4	Electronic Computer and Switching Systems Specialist
306X6	Secure Communication Systems Maintenance Specialist
361XX	Communications Antenna and Cable Systems Installation and Maintenance Specialist
391X0	Maintenance Data Systems Analysis Specialist
392X0	Maintenance Scheduling Specialist
452X4	Tactical Aircraft Maintenance Specialist
454X0A	Aerospace Propulsion Specialist, Jet Engines
454X1	Aerospace Ground Equipment Mechanic
454X6	Airlift Electrical and Environment Systems Specialist
455X1X	Avionics Guidance and Control Systems Specialist
455X2X	Communication and Navigation Systems Specialist
457X0	Strategic Aircraft Maintenance Specialist
457X2	Airlift Aircraft Maintenance Specialist
462X0	Aircraft Armament Systems Specialist
472X4	Vehicle Maintenance Control and Analysis Specialist
491X1	Communications Computer Systems Operator
491X2	Communications Computer Systems Programming Specialist
493X0	Communications Computer Systems Control Specialist
496X0	Communications Computer Systems Planning and Program Management Specialist
602X0	Passenger and Household Goods Specialist
602X1	Freight and Packaging Specialist
603X0	Vehicle Operator/Dispatcher
605X5	Air Transportation Specialist
645X0	Inventory Management Specialist
645X1	Materiel Storage and Distribution Specialist
645X2	Supply Systems Analysis Specialist
661X0	Logistics Plans Specialist
672X1	Financial Management Specialist
672X2	Financial Services Specialist
674X0	Financial Analysis Specialist
702X0	Administration Specialist
731X0	Personnel Systems Management Specialist

Table 1 (continued)
Specialties Included in the Present Study

AFSC	U.S. Air Force Job Title
732X0	Personnel Specialist
734X0	Social Actions Specialist
751X0	Education Specialist
751X1	Training Systems Specialist
811X0	Security Specialist
811X2	Law Enforcement Specialist
881X0	Paralegal Specialist
906X0	Medical Administrative Specialist

Transfer of Skills Questionnaire

Three different forms of the Transfer of Skills Questionnaire (TSQ) were devised: one for the retrainees, one for the supervisors and co-workers, and one for the non-retrainees. All forms were quite similar with the differences noted below.

For each of the 45 categories all respondents were asked to indicate: 1) whether the task was part of the job for a typical 5-level airman; 2) how much time is spent on that category for a typical 5-level airman; and 3) how many months it takes a newly assigned 3-level airman to become proficient, at a 5-level, on this category. In addition, for these 45 categories retrainees, supervisors, and co-workers were asked to indicate for the target airman (i.e., retrainee): 1) how many months it has taken to reach proficiency; and 2) the level of job.

Retrainees, supervisors, and co-workers were asked to indicate the target airman's overall performance by rating 1) technical proficiency; 2) interpersonal proficiency; 3) time it took to reach proficiency after reassignment; and 4) if the target airman had not reached proficiency, estimated time to reach proficiency.

Retrainees, supervisors, and co-workers were asked to assess the work unit climate of the retrainee's job. They were asked to rate a series of statements (32 for retrainees, 27 for supervisors) on a seven-point scale. These statements dealt with: 1) the learning environment; 2) supportive organizational structure and policies; 3) encouragement of new ideas and transfer of training; 4) goal setting; 5) performance constraints; 6) co-worker support; and 7) supervisory behaviors. Finally, retrainees were asked to respond to five statements on their attitudes about retraining and reassignment.

In addition to the TSQ, all respondents were also asked to provide background information. This information included, among others, age, race, sex, and time spent in present position.

Table 2

Task Taxonomy Category Definitions

Clerical (CLER): Tasks such as filing, preparing forms, answering telephones, typing reports, and proofreading. Operating office equipment such as computers, typewriters, calculators, duplicating machines. Processing information related to military regulations, federal or state laws, contracts, and legal documents.

Personnel (PERS): Processing data/information about individuals, such as employment applications, performance reviews, disciplinary reports, media releases, information, and social services.

Maintaining Inventories and Records (RECOR): Maintaining materials/merchandise/ supplies/equipment records. Ordering, receiving, maintaining, routing and accounting for inventory. Preparing, analyzing and maintaining records of financial dealings, property, assets.

Mathematical (MATH): Tasks such as adding, subtracting, multiplying, dividing. Computing statistics using formulas and equations. Locating statistics/data in graphs/tables/charts.

Physical/Manual Labor (MANUL): Nontechnical manual and physical tasks such as sweeping, lifting, cleaning, sawing, lubricating, drilling, cutting, hoisting, chipping, and planing.

Manufacturing/Fabricating (MANUF): Manufacturing things (e.g., pressing, mixing, forging, grinding) from materials such as sheet metal, metal tubing, brick, plastic, rubber, paper, or lumber.

Construction (CONST): Building/maintaining structures made of brick, stone, lumber, asphalt, or concrete, such as walls, floors, cabinets, houses, bridges, towers, roads, or runways.

Medical-Patient Care (MEDP): Interacting with patients. e.g., bandaging, giving injections, applying medicines, drawing blood; reading/interpreting medical charts, thermometer readings, test results.

Oral and Written Communication (COMM): Reading, speaking, writing, expressing ideas, such as in letters, books, reports, phone calls, orders, directions/instructions, and presentations.

Planning/Problem Solving (PROB): Using information to anticipate/figure out/solve problems, and plan steps and procedures required to reach a solution to the problem.

Science and Engineering (SCI): Using technical information such as aerial photos, weather forecasts, maps, engineering plans, blueprints, circuit diagrams, or information about people, events, places to test theories/products/equipment.

Artistic-Audio/Visual (ART): Tasks such as designing/producing photographs, movies, recordings, drawings; playing musical instruments, singing, dancing, acting; operating videotape players, slide projectors, etc.

Food Preparation (FOOD): Preparing/cooking food, using/producing menus, recipes, nutrition guides, food requests and estimates.

Fabric/Rope Work (ROPE): Sewing, stitching, threading, weaving, combining, or separating materials such as fabric, thread, rope, material, fiber, and string.

Managing Others (MANAGE): Supervising and evaluating others, e.g., setting goals, coordinating activities, assigning work, evaluating performance, conducting meetings, settling conflicts.

Training (TRAIN): Explaining ideas/procedures to others, demonstrating how a task is done, monitoring learner progress, providing feedback on mistakes, preparing lesson plans, course outlines, etc.

Surveillance (SURVL): Detecting and recognizing objects that are difficult to see, tracking and pursuing moving targets, using weapons, enforcing rules or laws.

Cryptographic Equipment Maintenance (CRYPT): Maintaining synchronizers and teletypewriter, cryptographic, automatic secure voice communications, fixed record communications terminal equipment.

Flight Control and Navigation Systems Maintenance (FCMAIN): Maintaining navigation and flight instruments, and inertial navigation, fuel savings advisory, automatic flight control, damping, quantity indicating and position indicating systems.

Munitions Handling (MUNI): Installing, testing, inspecting, maintaining, configuring, reconditioning, storing, transporting, and disposing of munitions or ordnance devices. Transfer of Skills Questionnaire

Table 2 (continued)

Task Taxonomy Category Definitions

Test, Measurement, and Diagnostic Equipment Maintenance (TEST): Maintaining and repairing equipment such as frequency generating and measuring, electronic precision measuring, waveform analyzing, electro-mechanical, optical precision measuring, and voltage, current, and impedance equipment.

Pneudraulic Systems Inspection and Maintenance (PNEUD): Installing, removing, inspecting, maintaining and performing operational checks of hydraulic, pneudraulic systems and components.

Computer Systems Maintenance (COMP): Maintaining, installing and modifying keyboards, printers, central processing units, magnetic tape units, video display equipment, input/output assemblies, etc.

Ground Based Communications Systems Maintenance (GBCOM): Maintaining receivers, transmitters, transceivers, antenna systems, ground radio and auxiliary equipment (Note: not telephone systems).

Aircraft Instrumentation Maintenance (INSTR): Preparing, installing, calibrating, aligning, removing, replacing, inspecting, testing, troubleshooting and maintaining installed aircraft instrumentation equipment.

Meteorological Equipment Maintenance (METE): Maintaining meteorological weather radar, nonelectronic meteorological instruments and solid state barometers, and wind, temperature, visibility, and cloud weather equipment.

Aircraft Armament Delivery and Control Systems Maintenance (ARM): Inspecting, maintaining, and performing operational or functional checks on aircraft installed suspension, launch, and release systems, gun systems, aircraft armament systems, and support equipment.

Take-off and Landing Control Systems Maintenance (TAKEF): Maintaining RAPCON, video mapper, landing control, and ASR, PAR, or OPS trailer systems.

Aircraft Navigation and Weapon Radar Maintenance (NAVI): Installing, removing and maintaining navigation, radar, digital modular avionic, and weapons release computer systems.

Malfunction and Recording Control Systems Maintenance (MALF): Maintaining malfunction analysis detection and recording systems, general purpose or navigational computers, and digital modular avionic systems.

Telephone Systems Maintenance (TELE): Installing, repairing and maintaining key telephone systems, wiring or components, special circuits, and mobile or fixed switching center equipment.

Aircraft Electrical Systems Maintenance (ELEC): Inspecting, maintaining and isolating malfunctions in aircraft electrical systems and circuit components.

Aerospace Ground Equipment Circuit Component Maintenance (CIRC): Maintaining heating and refrigeration systems and equipment coolers. Maintaining equipment enclosures, chassis, drives, and avionics aerospace ground equipment (AGE) such as AGE engines, motors and generators.

Air and Ground Based Navigational Aids Maintenance (AGNAV): Maintaining transponders and altimeters, and omnirange, beacon, instrument landing, long range navigation, and automatic direction finder systems.

Satellite, Missile, and Antenna Systems Inspection (SAT): Preparing, installing, and maintaining antenna, missile, and satellite instrumentation systems.

Avionics Test Equipment Maintenance (AVION): Maintaining displays and indicators, computer inertial, and radio frequency test stations, avionic system mockups, aircraft mockups and manually tested equipment.

Quality Control (QUAL): Performing quality control, quality assurance, or contract monitor functions.

Data Retrieval and Storage Systems Maintenance (DATA): Maintaining BRITE II, simulator PIDP systems, drum units and plotters (e.g., XY plotters).

Space-based Assets Control (SPACE): Maintaining space systems equipment. Performing maintenance/job control or sensor technician functions.

Propulsion and Engine Systems Maintenance (PROP): Removal, disassembly, inspection, repair, assembly, and installation of propulsion and engine systems and related equipment.

Data Editing

Editing data was done by making sure that the appropriate response was made for the first question to each category. If a particular category was appropriate to an AFS, then there should be a "1"; otherwise it should be left blank. If there was a value greater than one for response, it was changed to a 1. If that answer was left blank, and any of the remaining questions in that category were answered, then the answer was converted to a 1.

The necessity to clean up the data suggests some problems with study. Respondents may have misunderstood the instructions: 1) by putting an inappropriate value to the first question to a category; 2) by finding a category not appropriate to an AFS but still answering questions about it; or 3) by forgetting to respond to the first question in a category. It is also possible that participants in using the answer sheet misaligned their response with the appropriate part of the answer. One problem in analyzing these data is that it is not always possible to tell which of the above errors respondents made.

Outliers

Concern that extreme scores might influence the results led to the exclusion of outliers. For these data an outlier was defined as one whose included number of categories was more than four standard deviations from the mean. In the case of the retrainees the mean number of categories marked was 8.60 with a standard deviation of 4.16. Five cases exceeding 26 were dropped, leaving 929. With supervisors the mean number of categories was 8.30 with a standard deviation of 3.70. One case with 39 responses and two cases which matched retrainee outliers were dropped leaving 640. Among the co-workers the mean number of categories marked was 8.75 with a standard deviation of 3.67. One outlier with 27 responses was dropped as well as three cases who matched retrainee outliers, leaving 462. With respect to the non-retrainees the mean number of categories marked was 9.00 with a standard deviation of 4.54. Excluding responses in more than 27 categories led to dropping two, leaving 528.

RESULTS

Part of Job

One question in this investigation was which of the taxonomy categories are applicable to which AFS's. Counts of the number of respondents per AFS were recorded for each of the 40 categories. Certain categories appear to be part of most AFS's (especially, Clerical, Mathematical, Oral and Written Communication, Planning/Problem Solving, and Training) and certain categories rarely appeared in any AFS (especially, Medical-Patient Care, Food Preparation, Meteorological Equipment Maintenance,

Aircraft Armament Delivery and Control Systems Maintenance, and Take-off and Landing Control Systems Maintenance). Typically 8.5 to 9.0 categories were chosen per AFS, with most AFSs showing between 6 and 12 categories.

Evaluation

In their earlier taxonomy, Lance, Kavanagh and Gould (1995) proposed three criteria to evaluate its usefulness: 1) its face validity by examining the part-of-job ratings; 2) the reliability of the respondents and of the categories by examining relative time spent (RTS) ratings; and 3) the distinctiveness of the categories by examining the correlations and factor analysis of the RTS ratings. In the current investigating those criteria in examining the revised taxonomy with the retrainees' data will be used as well as other criteria.

Face Validity

The percentage of retrainees making part-of-job endorsements were examined. As can be seen in Table 3, the percentages were quite varied ranging from a low of 1% for Take-off and Landing Control Systems Maintenance to 89% for Clerical. The representative AFS's for each category are among those expected to have high part-of-job ratings. The percentages of part of job are also comparable to those in Lance, Kavanagh and Gould's (1995, Table 8) study.

Reliability

There are several ways to measure reliability in job analysis, but two methods are most prominent: using Pearson interrater reliability correlations and intraclass correlations. For interrater reliability raters are treated as variables and categories are treated as subjects. Correlations were then calculated for retrainees within the same AFS on the RTS ratings. The average correlation within AFS ranged from .34 to .80 with an overall average correlation of .61, comparable to Lance, Kavanagh and Gould (1995, p. 18). Using the Spearman-Brown correction for the average number of respondents per AFS (19.35) resulted in an $r = .97$.

Lahey, Downey, and Saal (1983) report a number of intraclass correlations (ICCs) which can be calculated, depending upon the type of reliability one wishes to examine. Lance, Kavanagh and Gould (1993, 1995) have used ICC (1, k) which indexes the reliability of the mean of k judges' ratings and is calculated by subtracting the within mean square from the between mean square (i.e., between-AFS mean square obtained in an analysis of variance) and dividing by the between mean square. Those ICCs and RTS means and standard deviations are presented in Table 4.

Table 3

Part-of-Job Ratings

Task Category	Part-of-Job Percentage	Representative AFS with High Percent Part of Job
CLER	89%	732X0 - Personnel Specialist
PERS	46%	732X0 - Personnel Specialist
RECOR	56%	304X4 - Ground Radio Communications Specialist
MATH	67%	674X0 - Financial Analysis Specialist
MANUL	52%	602X1 - Freight and Packaging Specialist
MANUF	6%	361XX - Communications Antenna & Cable Systems
MEDP	2%	906X0 - Medical Administrative Specialist
COMM	88%	751X1 - Training Systems Specialist
PROB	83%	491X2 - Communications Computer Systems Programming
SCI	28%	251X0 - Weather Specialist
ART	18%	201X0 - Intelligence Operations Specialist
FOOD	4%	114X0 - Aircraft Loadmaster
ROPE	4%	361XX - Communications Antenna & Cable Systems
MANAGE	62%	114X0 - Aircraft Loadmaster
TRAIN	76%	751X1 - Training Systems Specialist
SURVL	10%	272X0 - Air Traffic Control Operator
CRYPT	7%	455X2 - Avionic Communication Specialist
FCMAIN	8%	455X1X - Avionics Guidance & Control Systems Specialist
MUNI	4%	454X6 - Airlift Electrical & Environmental Systems Specialist
TEST	5%	304X4 - Ground Radio Communications Specialist
PNEUD	12%	457X0 - Strategic Aircraft Maintenance Specialist
COMP	15%	305X4 - Electronic Computer & Switch Systems Specialist
GBCOM	5%	304X4 - Ground Radio Communications Specialist
INSTR	6%	455X1X - Avionics Guidance & Control Systems Specialist
METE	2%	304X2 - Meteorological & Navigation Systems Specialist
ARM	2%	457X0 - Strategic Aircraft Maintenance Specialist
TAKEF	1%	304X4 - Ground Radio Communications Specialist
NAVI	3%	455X2X - Communications & Navigation Systems Specialist
MALF	4%	455X1X - Avionics Guidance & Control Systems Specialist
TELE	11%	361XX - Communications Antenna & Cable Systems
ELEC	11%	454X6 - Airlift Electrical & Environmental Systems Specialist
CIRC	3%	454X6 - Airlift Electrical & Environmental Systems Specialist
AGNAV	5%	455X2X - Communications & Navigation Systems Specialist
SAT	4%	304X4 - Ground Radio Communications Specialist
AVION	3%	455X1X - Avionics Guidance & Control Systems Specialist
QUAL	23%	251X0 - Weather Specialist
DATA	7%	731X0 - Personnel Systems Management Specialist
SPACE	7%	277X0 - Space Systems Operations Specialist
PROP	9%	452X4 - Tactical Aircraft Maintenance Specialist

Table 4

Descriptive Statistics and Intraclass Correlations
for Relative Time Spent Ratings for Retrainees

Task Category	ICC		
	Mean	S.D.	(1,k)
CLER	5.20	2.90	0.916
PERS	1.94	2.70	0.898
RECOR	2.47	2.75	0.773
MATH	3.14	2.95	0.912
MANUL	2.02	2.72	0.916
MANUF	0.21	1.02	0.754
CONST	0.13	0.83	0.781
MEDP	0.06	0.58	0.508
COMM	5.22	2.92	0.815
PROB	4.59	2.91	0.696
SCI	1.49	2.69	0.898
ART	0.62	1.61	0.813
FOOD	0.09	0.70	0.708
ROPE	0.10	0.77	0.809
MANAGE	2.77	2.77	0.536
TRAIN	3.82	2.90	0.746
SURVL	0.49	1.83	0.903
CRYPT	0.24	1.12	0.789
FCMAIN	0.34	1.41	0.955
MUNI	0.12	0.77	0.670
TEST	0.20	1.08	0.859
PNEUD	0.59	1.79	0.971
COMP	0.76	2.14	0.954
GBCOM	0.27	1.40	0.977
INSTR	0.30	1.33	0.943
METE	0.06	0.61	0.932
ARM	0.09	0.68	0.572
TAKEF	0.05	0.56	0.205
NAVI	0.16	0.99	0.935
MALF	0.18	1.02	0.894
TELE	0.43	1.63	0.870
ELEC	0.49	1.63	0.937
CIRC	0.12	0.89	0.641
AGNAV	0.18	1.04	0.920
SAT	0.14	0.88	0.569
AVION	0.14	0.91	0.849
QUAL	1.13	2.36	0.716
DATA	0.22	1.18	0.364
SPACE	0.22	1.20	0.498
PROP	0.37	1.41	0.918

The intraclass correlations are for the most part high. The exceptions are for categories that are infrequently part of job (Medical-Patient Care, Aircraft Armament Delivery and Control Systems Maintenance, Take-off and Landing Control Systems Maintenance, Data Retrieval and Storage Systems Maintenance, and Space-based Assets Control) and their low frequencies may be responsible for the low reliabilities. The other low reliability category is Managing Others , which also had reliability problems for Lance, Kavanagh, and Gould (1995, Table 9). This low reliability may be due to actual differences in managing others within an AFS or different interpretations of what managing others means.

Distinctiveness

Two approaches were used to measure the distinctiveness of the taxonomy's categories. First, the average inter-item correlation in the 40 x 40 matrix of correlations among the RTS ratings across AFS's was calculated. The correlation ($r = .06$) indicates that the categories are measuring distinctively different tasks.

Second, a principal components analysis with a varimax rotation was conducted of the RTS ratings. A scree plot of the eigenvalues suggested 7 factors accounting for 42 percent of the variance. These factors are interpretable and comparable to those obtained by Lance, Kavanagh, and Gould (1995, Table 10). More specifically, these factors can be interpreted as: I) instrumentation maintenance; II) general and administrative; III) aircraft electrical systems; IV) communications systems; V) interpretation of information; VI) physical labor; and VII) interaction with others. These results suggest the factors are distinct but fit into a broader categorization.

Agreement Between Retrainees, Co-workers, and Supervisors

One question raised was whether there was agreement between retrainees, co-workers, and supervisors on similar items. Correlations for each item on the TSQ were calculated between retrainees, co-workers, and supervisors . It was also possible that retrainees' responses might be less accurate due to their unfamiliarity with their current AFS. As a result, partial correlations were calculated between retrainees and co-workers and between retrainees and supervisors, controlling for the number of months retrainees were in their present AFS.

A few patterns are noticeable from these correlations: 1) partial correlations did not change the zero order correlations appreciably (in most cases, r 's did not change more than .01 in either direction); 2) correlations were not very strong between any group (r 's are usually less than .40 unless n's are small); and 3) correlations for later items in a category are often lower than for earlier items. It should be noted that responses left blank were treated as missing data.

A related question was whether estimates of time to proficiency were similar for retrainees, co-workers, and supervisors. Here responses left blank were treated as zeros. Analyses of variance for repeated measures were done for months to proficiency (MTP) were done for each of the 40 categories. If the value of F had a probability of less than .10, then t-tests for paired samples were performed. Those significant differences, when found, revealed that retrainees underestimated time to proficiency when compared to co-workers and supervisors.

Estimating Retraining Times

Lance, Kavanagh, and Gould (1993) have described a procedure to estimate cross-job retraining times. The procedure involves computing the mean MTP ratings for each of the categories. To estimate the retraining times between Job 1 and Job 2, comparisons are made between the mean MTP values for each of the categories. If one is considering movement from Job 1 (old job) to Job 2 (new job), differences between Job 1 and Job 2 values for which Job 2 values are higher are calculated and summed (differences for which Job 2 values are lower are set equal to zero). For movement from Job 2 to Job 1, the reverse is done. This allows asymmetric values for retraining times (i.e., it may be easier to move from Job 1 to Job 2 than vice versa).

Of the 929 retrainees included in the above analyses, 129 were dropped because MTP ratings could not be obtained for their old AFS from any of the samples. A further 61 retrainees are not included in the following analyses because they retrained from one specialty to another within the same AFS.

Lance (personal communication) provided the FORTRAN program to estimate retraining times, based on the typical airman. The program calculates, where possible, six values: 1) estimates based on comparisons between non-retrainees (old) and retrainees (new) (NNRTE); 2) estimates based on comparisons between non-retrainees (old) and supervisors (new) (NSRTE); 3) estimates based on comparisons between non-retrainees (old) and co-workers (new) (NPRTE); 4) estimates based on comparisons within retrainees' data (old and new) (RRRTE); 5) estimates based on comparisons within supervisors' data (old and new) (SSRTE); and 6) estimates based on comparisons within co-workers' data (old and new) (PPRTE).

In addition, three new variables were created by the summing the values for MTP for the target airman across the 40 categories without regard to old and new AFS for retrainees, supervisors, and co-workers (MTPR, MTPS, and MTPC, respectively). The resulting correlations are shown in Table 5. For certain patterns (i.e., old to new AFS's) there were more than one retrainee. This could result in the values for a pair of variables being repeated several times. Consequently, only one example of each possible change was included for correlations between pairs of Lance's six values (i.e., NNRTE, NSRTE, NPRTE, RRRTE, SSRTE, and PPRTE).

Table 5

Correlations between Cross-Job Retraining Time Estimates
and Summed Months to Proficiency

	NRRTE	NSRTE	NPRTE	RRRTE	SSRTE	PPRTE	MTPR	MTPS
NSRTE	.7864 (n = 268)							
NPRTE	.7960 (n = 264)	.7063 (n = 264)						
RRRTE	.8354 (n = 143)	.6310 (n = 143)	.6587 (n = 142)					
SSRTE	.6633 (n = 143)	.8659 (n = 143)	.6160 (n = 142)	.7313 (n = 155)				
PPRTE	.7510 (n = 142)	.6208 (n = 142)	.9239 (n = 142)	.6739 (n = 154)	.6511 (n = 154)			
MTPR	.3175 (n = 694)	.2609 (n = 694)	.2481 (n = 689)	.2743 (n = 371)	.2659 (n = 371)	.1864 (n = 369)		
MTPS	.2809 (n = 480)	.4088 (n = 480)	.2609 (n = 475)	.3224 (n = 269)	.4257 (n = 269)	.2080 (n = 267)	.2208 (n = 494)	
MTPC	.3898 (n = 352)	.3554 (n = 352)	.4770 (n = 351)	.2992 (n = 191)	.2550 (n = 191)	.4379 (n = 190)	.2716 (n = 359)	.4156 (n = 296)

Certain patterns are evident in the correlations, which are all significant. The nine variables can be classified in multiple ways. One classification would form groups by those involving MTP ratings (MTPR, MTPS, and MTPC), estimates involving non-retrainees (NRRTE, NSRTE, and NPRTE), and estimates within in a sample (RRRTE, SSRTE, and PPRTE). Another classification would form groups variables in those involving retrainees (MTPR, NRRTE, RRRTE), supervisors (MTPS, NSRTE, and SSRTE), and co-workers (MTPC, NSRTE, SSRTE). Generally correlations within a classification are higher than those across classifications. For example, NRRTE, NSRTE, and NPRTE are highly correlated, as are NSRTE, SSRTE, and MTPC.

Correlations with Cross-Job Retraining Times

Cross-job retraining times should be related to job similarity and job difficulty (Lance, Kavanagh & Gould, 1993; Truhon, 1993). Lance, Kavanagh and Gould (1993) found support for three hypotheses: 1) there should be a positive relationship between learning difficulty of the new job (NJLD) and cross-job retraining time estimates; 2) there should be a negative relationship between the learning difficulty of the old job (OJLD) and the cross-job retraining estimates; and 3) cross-job retraining should easier for jobs involving the same aptitude requirements (MAGEON).

For these analyses the data were based upon the 731 retrainees mentioned earlier. The learning difficulty of the old and new AFS were obtained from benchmarked values (cf., Mumford, Weeks, Harding, & Fleishman, 1987). The variable measuring similarity of aptitude requirements was created by indicating whether retraining was within (=0) or across (=1) MAGE aptitude area requirements. Some AFS's have requirements in two MAGE areas. If the old AFS had one of the two area requirements of the new AFS, then a value of .5 was possible.

The results of these hypotheses can be seen in Table 6. The first hypothesis is supported. Correlations between NJLD and cross-job retraining time estimates are all positive, significant, and close to the value reported by Lance, Kavanagh and Gould (1993). The second hypothesis was not supported. Only two of the correlations are significant and only one of those was in the predicted direction. The third hypothesis was not confirmed with correlations near zero in most cases.

Table 6

Correlations Between Learning Difficulty, Aptitude Requirements and Cross-Job Retraining Estimates

Index	NNRTE	NSRTE	NPRTE	RRRTE	SSRTE	PPRTE
NJLD	.3935*	.4087*	.3170*	.4985*	.4382*	.1443*
OJLD	.0902	.0378	.0195	.1971*	-.2027*	.0526
MAGEON	-.0070	-.0630	.0217	-.1541	.0579	.0472

*p < .01

Lance, Kavanagh & Gould (1993) also examined this third hypothesis by testing for a Old-MAGE x New-MAGE interaction effect in 4 x 4 analysis of variance. In the current study, the 739 retrainees were classified by their old AFS into four categories as well as for their new AFS (Those

AFS's with two area requirements were classified by their primary area). Multivariate analyses of variance were then performed because the results from Table 5 suggest that NRRTE, NSRTE, and NPRTE form one group and RRRTE, SSRTE, and PPRTE form another group.

The results of those analyses revealed significant main effects for the Old-MAGE and New-MAGE variables as well as a significant Old-MAGE x New-MAGE interaction. Examination of the cell means suggests that: 1) retraining from Administrative specialties is more than difficult than other specialties (partly comparable to Lance, Kavanagh and Gould's [1993] results); 2) retraining into mechanical and electronic jobs is more difficult than other specialties (comparable to their results); and 3) with the exception to Electronic area, retraining times with MAGE areas were lower than across areas (partly comparable to their results).

CONCLUSIONS

The revised taxonomy that Lance has developed (see Table 2) adequately covers the types of tasks seen AFS's. Using the criteria they have established, this taxonomy has face validity, reliability, and distinctiveness. It provides good estimates of cross-job retraining times which can support research in this area.

However there are also problems. The above taxonomy is meant to be an improvement from Lance, Kavanagh and Gould's (1995) taxonomy. The above analyses suggest that the revised taxonomy is as good as, but not necessarily better than, the earlier taxonomy. Some of the infrequently used categories have problems in reliability. However, this may be due to the AFS's that retrainees moved into rather than the adequacy of the taxonomy.

There may also be problems with the size of the TSQ. Asking more than 200 questions about specific aspects of a specialty may be overwhelming for respondents. This may explain the low return rate. It may be best for the taxonomy to be shortened, especially with the infrequently occurring categories. This may improve the overall quality of the taxonomy and may make the questionnaire less intimidating to respondents.

References

Ballentine, R.D. & Cunningham, J.W. (1981). Development of the General Work Inventory. Proceedings of the Annual Conference of the Military Testing Association, 1, 125-133.

Bell, J. & Thomasson, M. (1984). Job categorization project. Randolph AFB, TX: Occupational Analysis Program, United States Air Force Occupational Measurement Center.

Fine, S.A. (1957a). A reexamination of "transferability of skills" - Part I. Monthly Labor Review, 80, 803-810.

Fine, S.A. (1957b). A reexamination of "transferability of skills" - Part II. Monthly Labor Review, 80, 938-948.

Gordon, M.E., & Fitzgibbons, W.J. (1982). Empirical test of the validity of seniority as a factor in staffing decisions. Journal of Applied Psychology, 67, 311-319.

Lahey, M.A., Downey, R.G., & Saal, F.E. (1983). Intraclass correlations: There's more there than meets the eye. Psychological Bulletin, 93, 586-595.

Lance, C.E., Kavanagh, M.J., & Gould, R.B. (1993). Development and convergent validation of a methodology for estimating cross-job retraining times. Journal of Business and Psychology, 8, 67-90.

Lance, C.E., Kavanagh, M.J. & Gould, R.B. (1995). Development and evaluation of a task taxonomy to support research on cross-job transferability of skills (AL/HR-TP-1995-0003). Brooks AFB, TX: Armstrong Laboratory, Human Resources Directorate, Manpower and Personnel Research Division.

Lance, C.E., Mayfield, D.L., & Gould, R.B. (1993) A postdictive validation of cross-job retraining time estimates for 43 U.S. Air Force enlisted specialties. Military Psychology, 5, 173-186.

Lance, C.E., Mayfield, D.L., Gould, R.B., & Lynskey, M.C. (1991). Global versus decomposed estimates of cross-job retraining time. Human Performance, 4, 71-88.

Mumford, M.D., Weeks, J.L., Harding, F.D., & Fleishman, E.A. (1987). Measuring occupational difficulty: A construct validation against training criteria. Journal of Applied Psychology, 72, 578-587.

Ruck, H.W. (1986). Skill/knowledge commonalities in selected electronic specialties (AFHRL-TP-86-20). Brooks AFB, TX: Human Resources Laboratory, Manpower and Personnel Division.

Sparrow, J.A. (1989). The measurement of job profile similarity for the prediction of transfer of learning: A research note. Journal of Occupational Psychology, 62, 337-341.

Truhon, S.A. (1993). Predicting ease of movement between Air Force specialties. Unpublished paper.

**A STUDY OF THE FEASIBILITY OF AUTOMATING METHOD APC 44
FOR THE DETERMINATION OF TOTAL GLYCOLS IN WATER**

**Sailaja Vallabha
Lecturer
Department of Natural Sciences**

**Fayetteville State University
1200 Murchison Road, Newbold Station
Fayetteville NC 28305**

**Final Report for:
Summer Faculty Research Program
Armstrong Laboratory**

**Sponsored by:
Air Force of Scientific Research
Bolling Air Force Base, DC**

and

Armstrong Laboratory

August 1995

**FEASIBILITY OF AUTOMATING METHOD APC 44
FOR THE DETERMINATION OF TOTAL GLYCOLS IN WATER**

Sailaja Vallabha
Lecturer
Department of Natural Sciences
Fayetteville State University

Abstract

This study investigates the feasibility of automating the EPA method APC 44, used for the determination of total glycols in water samples. The study included investigating the importance of the various steps involved in the analysis and the reagents used in the analysis. It was crucial to understand the chemistry involved in following the sequence of steps for the addition of the reagents in this analysis. Automation of this method will drastically reduce the amount of time spent by an analyst in the sample preparation and the analysis thus reducing the cost. Moreover it would eliminate the waste generated, cut back the amount of reagents used and the glassware needed. Thus the analysis when automated will be cheaper, faster, safer, easier and most important will be environment friendly.

The APC 44 method determines the total amount of glycols present in the water sample by first converting the hydroxy-methyl groups of the glycols to formaldehyde by the periodate scission. In the next step any unreacted and excess of periodate is quenched by the sodium bisulfite. This is followed by the use of chromotropic acid method which forms a purple colored complex with the formaldehyde. Spectroscopy is used to colorimetrically determine the amount of total formaldehyde present in the samples. Any interfering formaldehyde is determined separately by omitting the oxidation step and by carrying out the chromotropic step alone. The difference between these two values will be indicative of the glycol that was converted to formaldehyde.

**FEASIBILITY OF AUTOMATING NY APC 44
FOR THE DETERMINATION OF TOTAL GLYCOLS IN WATER**

Sailaja Vallabha

Introduction

The NY APC-44 method is for the determination of total glycol content in water samples. The water samples analyzed in this lab come from air force base storm water runoff. The major sources of contamination are air conditioning coolants and automobile anti-freezes and flight lines where de-icing chemicals are used. The non potable water samples are passed through an ion exchange column to remove interferences such as nitrates. Nitrates cause a yellow color that may reduce the absorbance intensity of the sample. The sample is then oxidized to formaldehyde during a reaction with potassium meta periodate. After the addition of the periodate the sample is placed in a 60°C water bath for about 15 minutes to accelerate the hydroxy-methyl scission. Any unreacted periodate is removed by adding 500 ul of sodium bisulfite to the solution. The resulting formaldehyde then is colorimetrically determined by its reaction with chromotropic acid. Formaldehyde in water reacts with chromotropic acid and sulfuric acid to form a monocationic chromogen. The absorbance of this purple colored solution is read on a spectrophotometer at 580 nm and is proportional to the amount of the formaldehyde in the solution. Formaldehyde already present in the sample can be calculated by treating the sample with chromotropic and sulfuric acids, while omitting the oxidation step. There are no other major interferences other than nitrates and formaldehyde in this analysis.

The study initiated in this lab was to explore the feasibility of automating this method. Automation of APC-44 will result in quicker analysis and preparation on a microscale. The latter is very attractive due to the fact that the amount of waste generated is reduced by 75%. The Pesticides and Trace Organics lab that where I worked during the summer is constantly striving to develop new methodologies and to modify existing methods to be environmentally friendly. A switch to microscale techniques eliminates large scale waste. Once tested and implemented this method will prove to be cost effective and result in a faster analysis time.

Experimental Procedure

In order to automate the process of APC-44 it was essential to understand the individual and the collective roles of the reagents used in the method. The following discussion describes the method currently used for glycol analysis.

Apparatus Required

1. Spectrophotometer capable of measuring the absorbance (at 580 nm) of the chromogen produced by the reaction between formaldehyde and chromotropic acid
2. Magnetic stirrer and stir bar
3. Class A serialized Volumetric flasks of different volume: 10, 100, and 200 ml
4. Pyrex test tubes (30 ml volume) and caps
5. Test tube racks to hold the tubes and ion exchange columns
6. Volumetric pipets of various volumes (100 μ l, 0.5 ml, 1 ml, 2 ml, 5 ml, 10 ml and 20 ml)
7. Pyrex disposable transfer pipets (10 ml)
8. Bio-Rad 20 ml ion exchange chromatography columns

Reagents Needed

1. 1% Chromotropic acid reagent: Prepared by dissolving 0.10g of 4,5- dihydroxy 2,7-naphthalene disulfonic acid sodium salt (Eastman Kodak Company, Rochester, New York). This solution needs to be prepared fresh on the day of the analysis
2. Concentrated sulfuric acid
3. Millipore water obtained from a Milli - Q water system
4. Ethylene glycol stock solution: 1.0g of neat standard ethylene glycol is dissolved in millipore water and made up to volume in a 100 ml volumetric flask. The stock solution can be used upto six months
5. Ethylene Glycol Intermediate standard (100 μ g/ml) : Diluting about 1.0ml of the stock solution to 100 ml. This intermediate standard has to be prepared on the day of the analysis. A series of working standards are prepared by serial dilution of the intermediate standard to obtain concentrations from 0.05 μ g/ml to 3.0 μ g/ml of ethylene glycol. All working standards are made with millipore water
6. 2% Sodium bisulfite solution: Weigh out 2 g of sodium bisulfite and dissolve in a 100 ml flask using millipore water. The solution is stable for 30 days
7. Potassium meta periodate solution (0.01M): Dissolve about 0.23 g of reagent grade potassium meta-periodate in a 100 ml flask with millipore water. Use a magnetic stir bar to accelerate the dissolution process. The solution is stable for thirty days
8. Dowex MR-3 mixed bed resin, nuclear grade Fisher Scientific: A medium porous mixture of strong acid cation exchanger in hydrogen form and strong base anion exchanger in hydroxyl form. Add 150 ml of resin to an erlenmeyer flask and 200 ml of millipore water to form a slurry. The is swirled and the liquid is decanted. This process is repeated three times. The resin can be stored with about 100 ml of millipore water in a covered erlenmeyer flask

9. Formaldehyde stock solution (1.0 mg/ml): Dilute 2.7 ml of 37% formalin neat standard in a liter of millipore water. This solution is standardized by titrating against sodium sulfite solution using Thymolphthalein as a color indicator. The intermediate standard is prepared by diluting 1.0 ml of stock solution to 100 ml with millipore water. The working standards are prepared by serial dilution of intermediate standard to obtain concentrations from 0.05 ug/ml to 2.0 ug/ml.

Ion exchange column preparation

Interferences found in the sample matrix are eliminated by passing the samples through an ion exchange column. The glycol and formaldehyde standards, Quality Control (QC) samples, a duplicate and a water blank are all treated in the same manner as the samples. The plastic ion exchange chromatography columns are placed in a test tube rack. The resin that is stored with millipore water is triple rinsed. After triple rinsing the resin add 200 ml of water and a magnetic stir bar. The contents of the flask are stirred to ensure uniform resin suspension. While stirring, 5 to 10 ml aliquots of the resin are transferred to the chromatography columns using a pyrex transfer pipet with its tip cut off. One column for each of the sample, standards, QC samples, duplicates and blanks.

Sample Preparation

A 10 ml volume of millipore water is placed on the ion exchange column and the eluant is discarded. A total of 30 ml of each sample, standards, QC samples, duplicates and the blanks and the water blank are run through the columns. Discard the first 10 ml aliquot into a waste receptacle. The second and third aliquots are collected in labeled test tubes. The samples are ready to be analyzed.

Sample analysis

Take 3.0 ml from each of the prepared samples and transfer to a new test tube. To each of these add 100 ul of the potassium meta periodate and mix. Place the test tubes in a 60°C hot water bath for 15 minutes. After 15 minutes remove the test tubes and add 500 ul of sodium bisulfite, mix well and wait for 5 minutes. After the five minutes add 100 ul of chromotropic acid and 6 ml of concentrated sulfuric acid, mix well. Allow to cool to room temperature. The glycols are oxidized to formaldehyde which in turn react with the chromophore to give a purple colored solution. The absorbance of the colored solution is read at 580 nm with a spectrophotometer. A calibration curve is generated by plotting the absorbance of the standards versus their concentrations. The total glycol present in the sample is read in ug/ml from the standard curve.

To analyze for formaldehyde take 4.0 ml of prepared sample and transfer to a test tubes. The analysis of formaldehyde is simply the addition of the chromotropic acid and 6 ml of sulfuric acid to the 4.0 ml aliquot. The absorbance of the standards is plotted against their concentration and the formaldehyde content of the

samples in (ug/ml) is determined from the calibration curve. The difference between (i) the amount of total glycol present in the sample and (ii) the amount of formaldehyde, will be a measure of the amount of glycols present in the sample.

Modifications to NY APC 44

High Pressure Liquid Chromatography (HPLC) with fluorescence detection was the instrument for automating this process. The lab is equipped with 4 Hewlett-Packard 1050 and 4 Hewlett-Packard 1090 HPLCs. They are equipped with diode array and fluorescence detectors. The next objective was to determine the exact requirements and hardware needed to automate NY APC-44.

The best choice of columns for this analysis was narrowed down to either a nitrile column or a diol column.

The 1050 HPLC units are equipped with post column reaction units that are binary pumping units. To determine whether this pumping system would be adequate or not for the automation, a series of analyses were performed. The different permutations would enable us to determine if all of the reagents (periodate, sodium bisulfite, chromotropic acid and sulfuric acid) would need to be pumped individually on to the column or if they could be combined in any way. The normal sequence of events and reagent addition to the sample are:

- (i) addition of potassium meta-periodate
- (ii) samples placed in hot water bath for 15 minutes
- (iii) addition of sodium bisulfite
- (iv) 5 minute wait period
- (v) addition of chromotropic acid
- (vi) addition of sulfuric acid

The various modifications and conditions tried are listed below:

- Control: the procedure followed as per method APC 44
- Condition 1: instead of adding individually a mixture of chromotropic and sulfuric acid added to sample
- Condition 2: The wait time is reduced to 2 minutes instead of 5 minutes following the addition of sodium bisulfite
- Condition 3: The 5 minute wait period was omitted
- Condition 4: Periodate and sodium bisulfite were combined together and added to the sample
- Condition 5: Sodium bisulfite eliminated from the process

Results from the modifications

The results from condition 1 are shown in figure 1. The results do not deviate much from the control analysis. This proves that the analysis could be carried out using a mixture of the chromotropic and sulfuric acids with good results. A single pump can be used to add these two reagents to the sample without compromising the results.

There is a 5 minute wait period between quenching the oxidizing agent by the bisulfite and the addition of the chromophore. Condition 2 explores the need for the 5 minute wait period. In Figure 2 where the results from this condition were plotted against the control indicate that there is no deviation from the regular analysis. To further confirm that the wait period may not be essential for the final result the wait period was eliminated in Condition 3. The results are shown in figure 3, which indicate the wait period is not needed in the sequence of operation.

The results from Condition 4 as expected show that it is impossible to get any absorbance readings from the samples when sodium bisulfite was combined with the periodate. The periodate was not allowed to convert the glycols to formaldehyde before it was neutralized by the bisulfite. The results of this modification plotted against the control is shown in figure 4. The results from all these modifications are shown in figure 5.

In condition 5 the sodium bisulfite was omitted. The final solutions were all yellow in color unlike the expected purple color. The absorbance values plotted against the concentration are shown in figure 6, emphasize the need for the bisulfite addition.

Conclusion

In summary, automating the NY APC-44 method, can reduce a six step process to a four step process. However, at this point all of the reagents currently used seem necessary. The meta-periodate is needed to perform the hydroxy methyl scission, sodium bisulfite to remove the excess periodate (so it will not interfere with the development of the chromogen), chromotropic and sulfuric acids to form the colored complex that can be detected. The procedure to be used in the automated analysis will involve only four steps. The suggested sequence of events for the automated method are as follows:

- (i) addition of potassium meta-periodate
- (ii) sample heated while on the column
- (iii) addition of bisulfite
- (iv) addition of a mixture of chromotropic and sulfuric acids

The mobile phase would consist of water and periodate which oxidizes the sample. The sample injected on to the column will react with the potassium meta-periodate.

The HPLC 1050s equipped with column heaters and post-column reaction heaters eliminate the water bath completely. Since the second step in the automated process is carried out simultaneously with the sample elution this further reduces the analysis time.

As a post column step, sodium bisulfite can be added to the sample. Since our investigations have shown that there is no need to wait between steps iii and iv, a mixture of chromotropic and sulfuric acids could also be added as a post column step.

Preliminary results have shown that a 50 ul injection of the low concentration glycol standard can be detected using a fluorescence detector. This small sample size suggests that the analysis time will be greatly reduced. Thus the HPLC systems would have to be equipped with tertiary pumping units in order to add all the reagents to the sample. This is a small initial expense that would be recovered within a very short time. At the present time the APC-44 method requires about two days for the sample preparation and the analysis. By automating this method the analysis time can be reduced to about half a day which is very attractive. The small sample size and simultaneous addition of the reagents reduces the analysis time by 75 to 80%. The analysis time is approximated to be about 3 to 4 minutes based on our observations. The sample preparation now would be microscale which is the way the Pesticides/Trace Organics lab at AL/OEA is headed towards. Once the lab acquires the required hardware and equipment, the actual parameters for automation can be determined and the method will be much cleaner, safer (as it eliminates handling of the corrosive sulfuric acid), easier (due to decreased amount of sample preparation) and most importantly reduces the time an analyst has to spend performing the analysis.

Recommendations

Once the hardware is in place, all of the modifications need to be repeated to ensure their accuracy. Experimental conditions such as flow rates and column temperatures need to be optimized. The analysis time may be further reduced if the oxidation step can be accelerated by heating the sample on the column. The flow rate will have to be based on the the time required for the heating of the sample. Another modification that would have to be investigated is combining steps ii, iii and iv. The combined addition of sodium bisulfite, chromotropic and sulfuric acids would certainly be most attractive since this would eliminate the purchase of new pumping systems and the process will be further simplified.

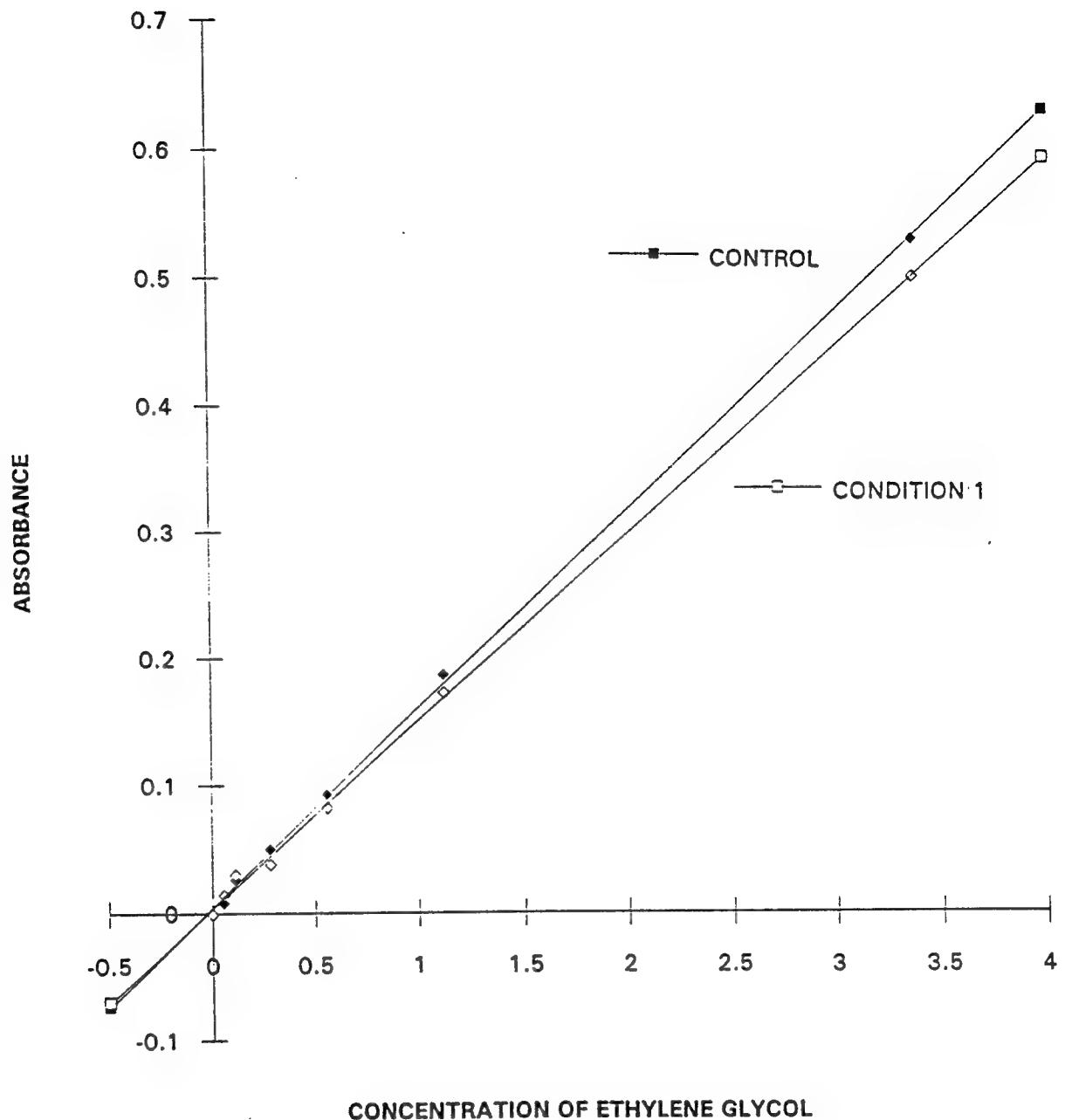


Figure 1. Plot of Control versus Condition 1

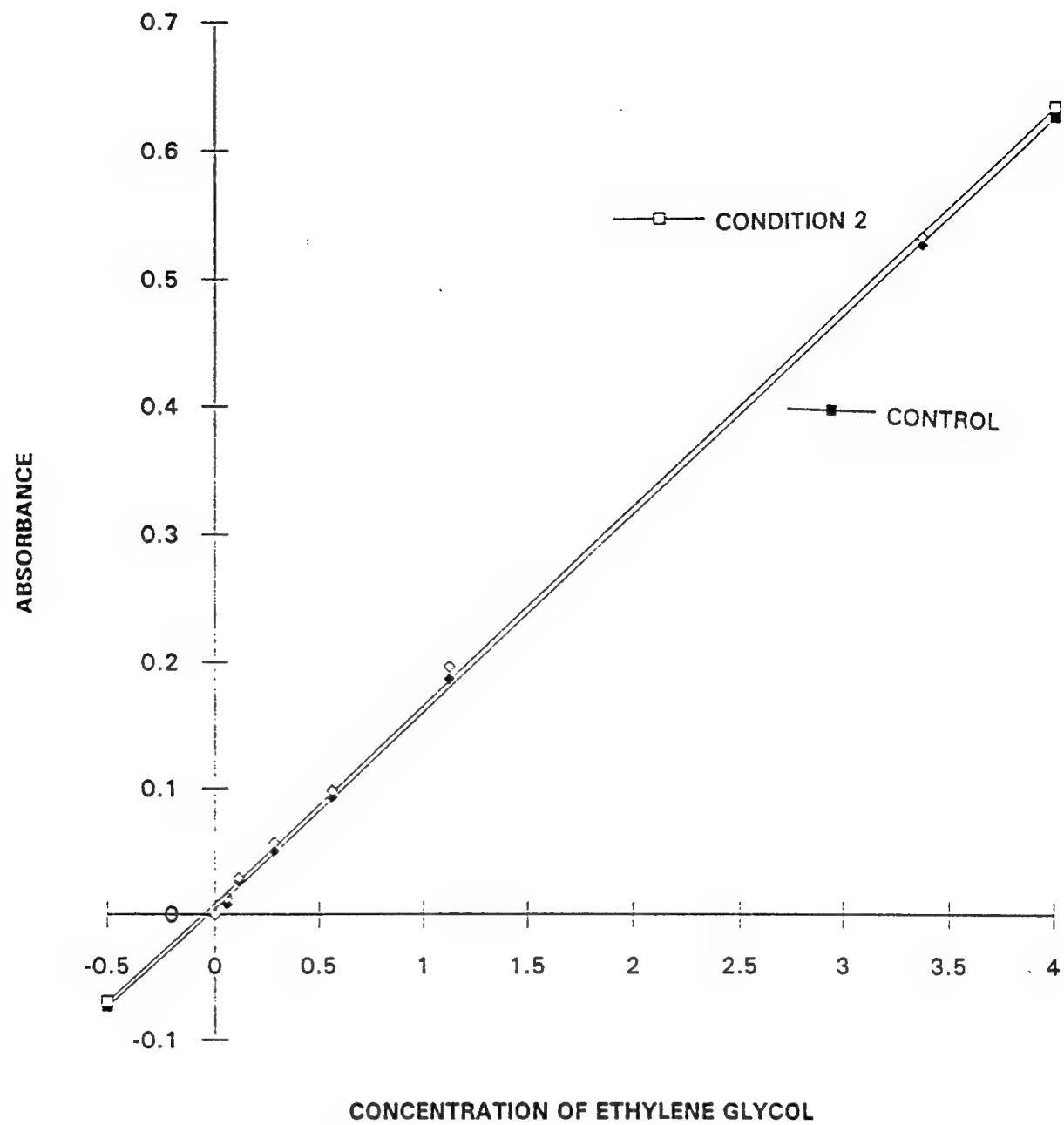


Figure 2. Plot of Control versus Condition 2

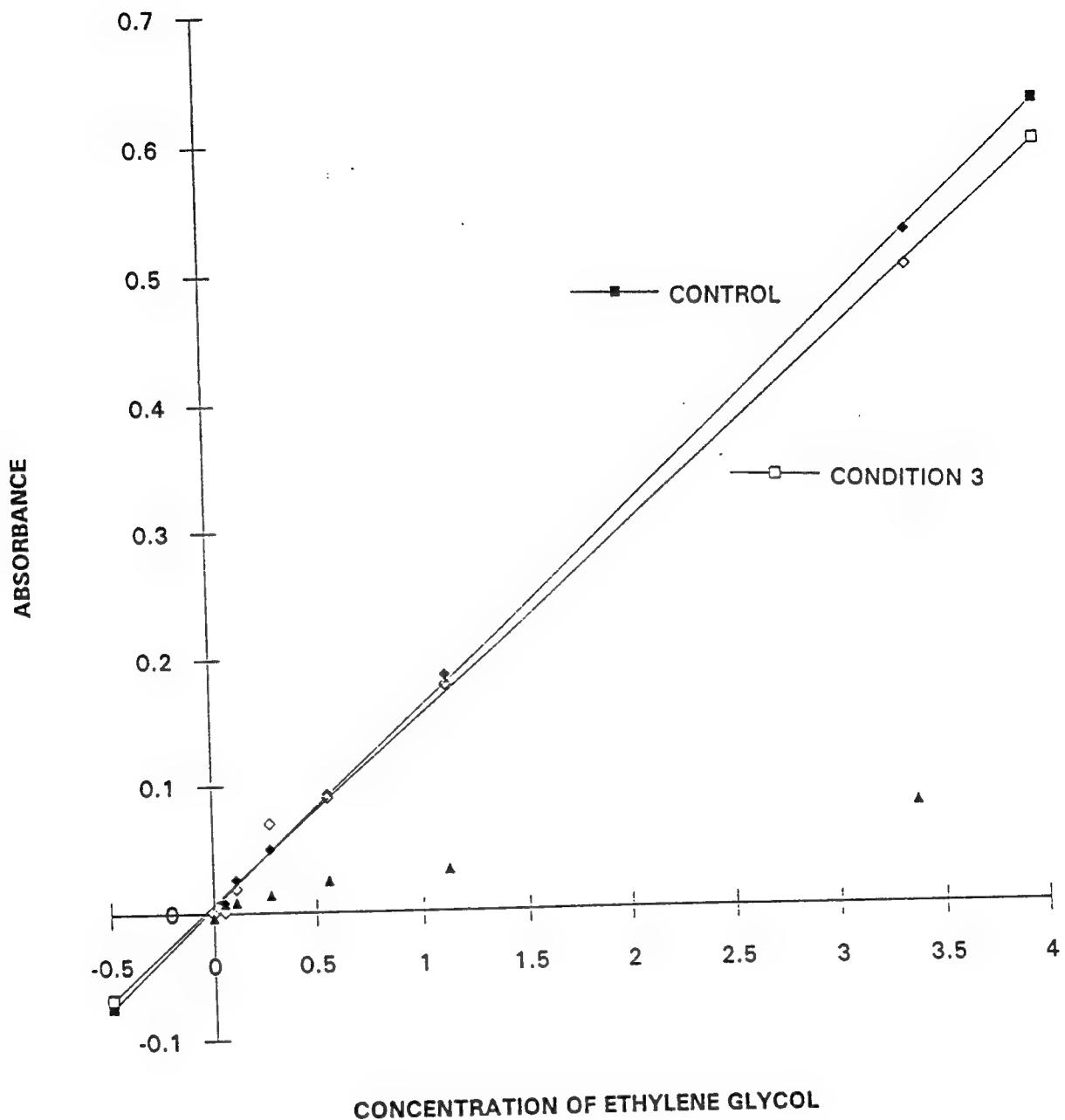


Figure 3. Plot of Control versus Condition 3

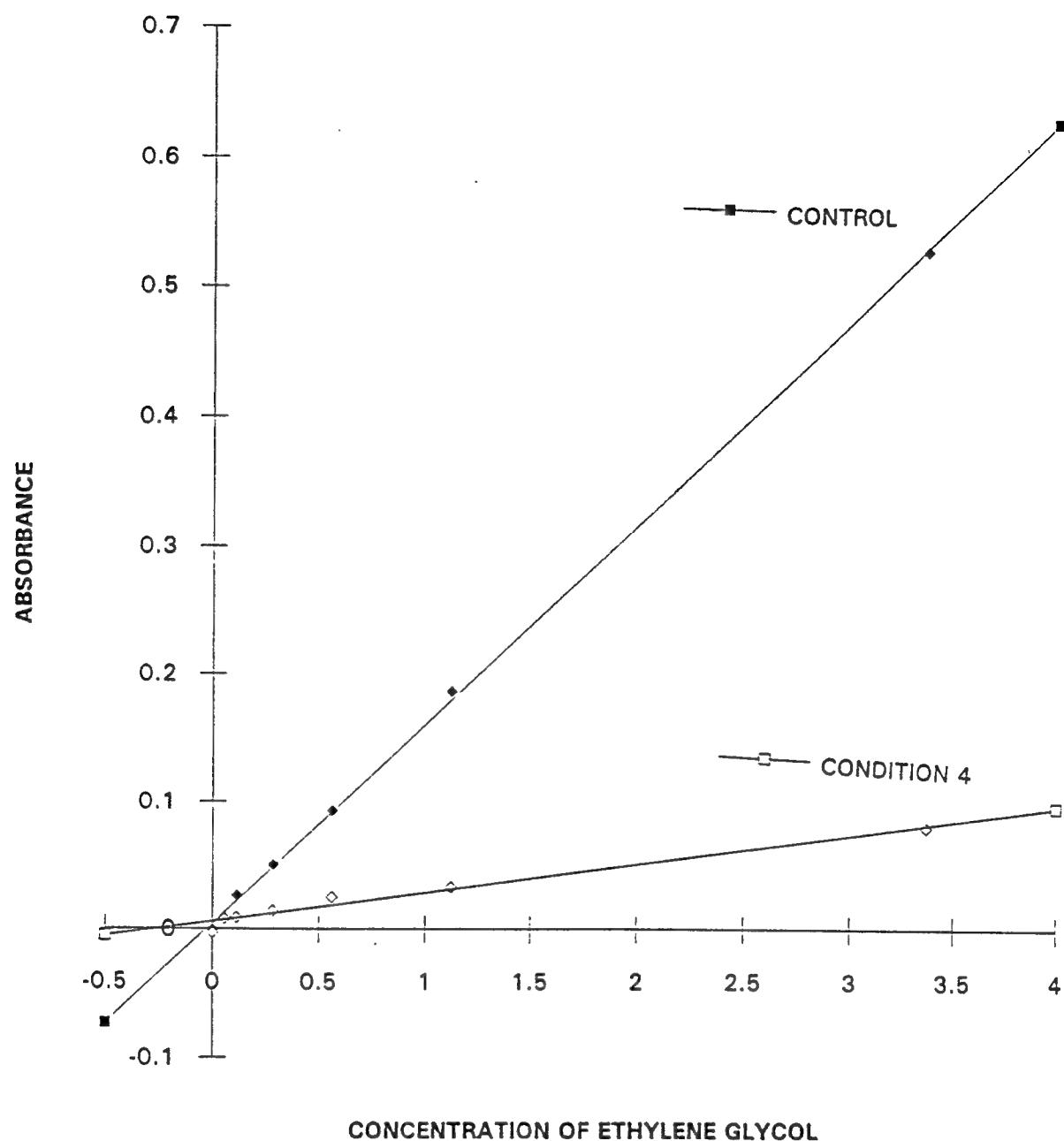


Figure 4. Plot of Control versus Condition 4

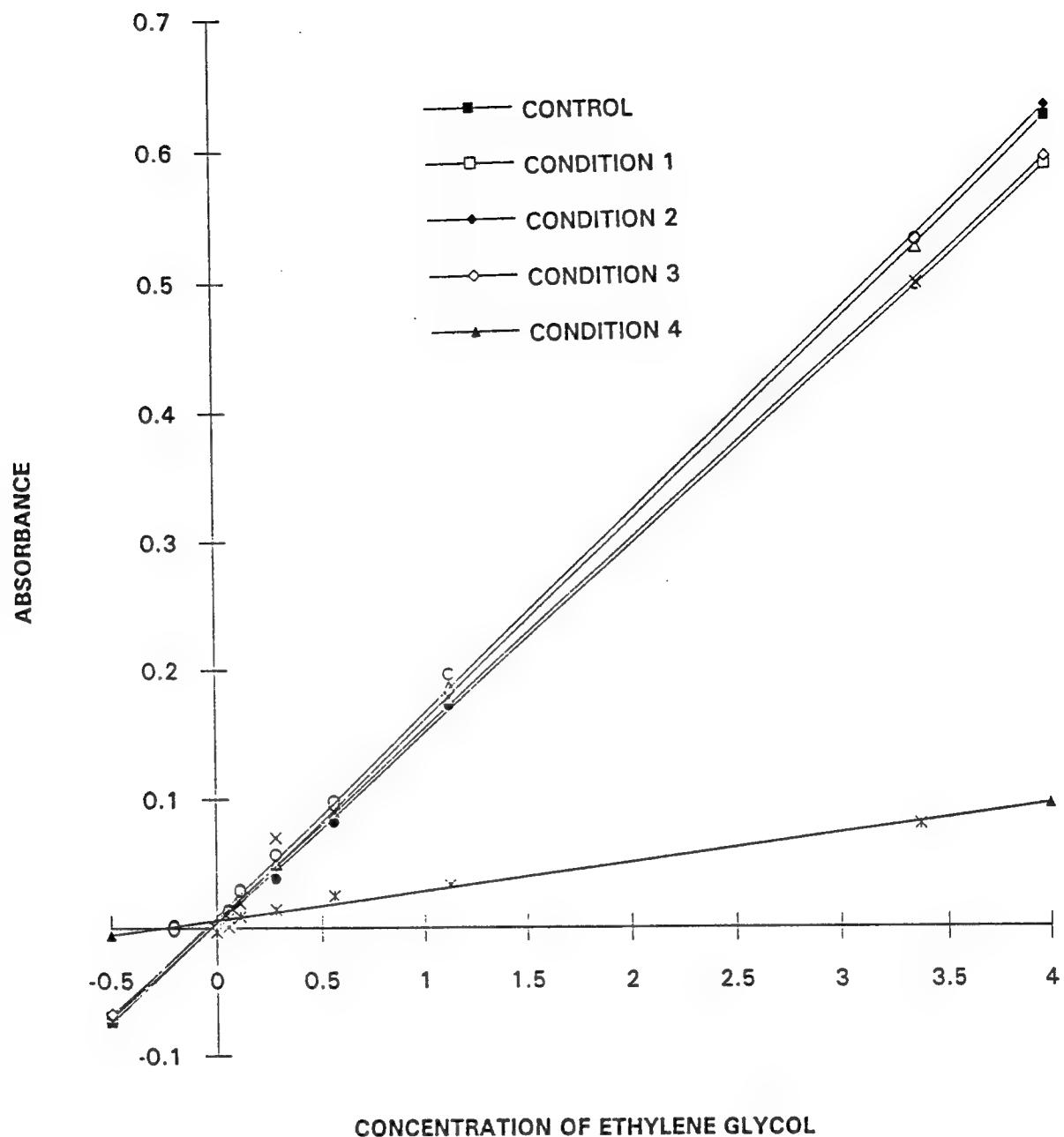


Figure 5. Plot of Control versus the Conditions 1 - 4

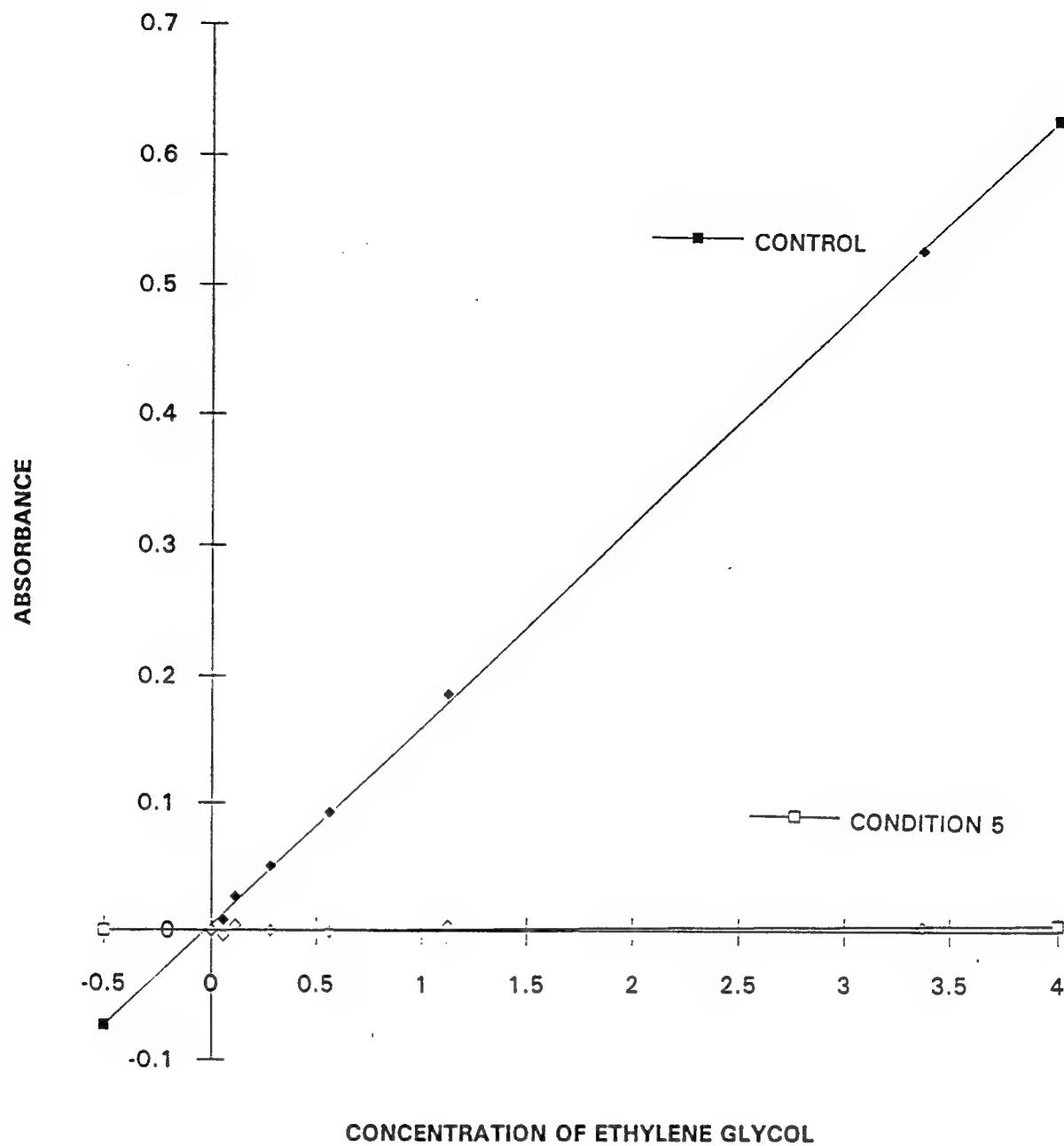


Figure 6. Plot of Control versus Condition 5

Optical Sensor Research in the Environics Sensors Program

Brian S. Vogt
Professor
Department of Chemistry

Bob Jones University
Greenville, SC 29614-0001

Terence Tipton
Research Chemist
Environics Directorate

Bruce J. Nielsen
Environmental Engineer, Program Manager
Environics Directorate

Armstrong Laboratory
Tyndall Air Force Base, FL 32403-6001

Final Report for:
Summer Faculty Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.

and

Armstrong Laboratory

July, 1995

Optical Sensor Research in the Environics Sensors Program

Brian S. Vogt
Professor
Department of Chemistry

Abstract

The purpose and focus of the Environics Sensors Program at Armstrong Laboratory (AL/EQ) are described. Previous research efforts have resulted in a uniquely designed laser probe that has been interfaced to a column and used to monitor the flow of contaminant plumes through sand packed into the column. Several improvements have been made to the laser-spectrometer sensor system and the laser probe. Work in progress centers on obtaining soil-column data to test a mathematical model developed by the Fate and Transport Group. This model describes the interactions of contaminants with layered soils. A full report was filed on base that details the accomplishments and potential research directions of the Sensors Group.

Optical Sensor Research in the Environics Sensors Program

Brian S. Vogt

Introduction

Contamination of soil components and ground water by fuels and solvents at Air Force sites is a matter of ongoing concern. Site characterization, remediation, and subsequent site monitoring to validate the remediation effort continue to be expensive but necessary efforts in eliminating such contamination. Both accurate site characterization and successful validation of remediation efforts are contingent on the availability of tools capable of measuring pollutants in a variety of settings. The Sensors Program at the Environics Directorate of Armstrong Laboratory (AL/EQ) is a research and development effort designed to help meet these Air Force environmental sensing needs.

In the broadest sense of the term, one can consider a sensor to be any device that interrogates part of the universe and responds by giving information about that locale. By this definition one can consider such diverse items as microscopes, bathroom scales, and the human ear to be sensors. They provide information that is optical, corporeal, or auditory in nature, respectively. In the context of environmental monitoring, however, a more restrictive description is appropriate: a sensor is a device that interrogates the environment and gives a signal related to a substance or parameter of interest. A sensor system consists of a sensor in combination with any electronics or other external devices necessary for the sensor to function properly. A combination pH electrode connected to a battery-powered meter is an example of a sensor system that meets these criteria. Although the electrode is the actual sensor, it cannot provide a useful signal without being connected to the circuitry of the meter. When the electrode is immersed in an aqueous medium, it gives an electrical signal related to the concentration of hydrogen ions (H^+) in that medium. The meter also functions to convert the signal to a more useful quantity (pH units) and displays it to the operator. Some sensor systems are capable of unattended operation for some given period of time--this distinguishes them from instruments used for

conventional laboratory analysis [1].

Environmental sensors may be classified conveniently as either optical or nonoptical in nature. Optical sensors rely on the interaction of light (electromagnetic radiation) with matter to extract information from the sample being interrogated. A change in light intensity as a result of absorption, emission, scattering, polarization, or refraction is monitored and converted into quantifiable electrical signals by an optical detection and transduction scheme. Most optical sensors discriminate between different analytes on the basis of the wavelengths of light they absorb or emit. The actual light signal is carried to and from the sample by inert optical components such as mirrors or fiber optics. In contrast, nonoptical sensors are intrinsically electrical in nature. Many are electrochemical devices, which oxidize or reduce the analyte by applying voltage to the surface of an electrode. Electrochemical sensors can discriminate between analytes on the basis of the voltage or current required to oxidize or reduce them.

The sensors group at AL/EQ has chosen to focus its efforts on developing fiber-optic sensor systems. The reasons for doing so are quite compelling [2-8]:

- Optical fibers have extremely high information-carrying capacity.
- The absence of electrical components frees them from interference from electromagnetic fields and problems due to short-circuiting.
- Unlike electrochemical sensor systems, all of which require reference electrodes, optical sensor systems require no external reference when making measurements.
- Fiber-optic sensors are easily miniaturized due to the small size of optical fibers.
- Optical fibers, especially the silica-clad silica variety, are highly resistant to corrosion.
- Their geometric flexibility makes them easy to deploy in a variety of environments.
- Sensor cost is potentially very low due to the low cost of fiber optics.
- System cost can be minimized by multiplexing many sensors to one optical detection apparatus.

One area of wide-spread interest is the use of optical fibers for direct spectroscopic sensing. In this approach, optical fibers carry probing light

to the sample of interest and transmit light from the sample back to the optical system for detection and analysis. While there are many spectroscopic methods that can be coupled to optical fibers, we have chosen to use laser-induced fluorescence (LIF) as the mainstay of our fiber-optic direct sensing efforts. The most common means of inducing fluorescence is to irradiate a sample with ultraviolet (UV) or visible light. Absorption of UV-visible light moves an electron from a given molecular orbital to one of higher energy. Energy emitted in the form of light when the electron returns to the lower-energy orbital is called fluorescence. Several other photophysical processes besides fluorescence can return the molecule to the lower energy state, but they are not relevant here and hence will not be discussed. Our work to date has involved substances requiring UV light for excitation. Because optical fibers attenuate UV light significantly, we use a laser to provide intense UV so that ample light reaches the sample even though some attenuation takes place.

Fluorescence analysis is particularly useful in direct spectroscopic sensing primarily for two reasons. First, it is an extremely sensitive method blessed with detection limits that few rivals can approach [9]. Under optimal experimental conditions it is even able to detect single molecules [10]. Being able to detect low concentrations of slightly-soluble components is of great importance when monitoring contaminated ground water. Second, fluorescence signals can provide information simultaneously in several dimensions, including intensity, wavelength, time, and polarization [11]. Information about the number of components and their concentrations in a sample is intrinsic to fluorescence signals. Spectral discrimination can be achieved by judicious choice of excitation and emission wavelengths because different substances absorb and emit different wavelengths of light. Temporal differences in the fluorescence signals of different substances provides another means of discrimination. It is natural to marry a high-information spectroscopic technique like fluorescence to high-information-capacity optical fibers. The combination of high sensitivity and information content unique to fluorescence analysis continues to stimulate researchers to use it in many different research settings [9].

Apparatus

Essential details of the apparatus and its application to soil-column experiments have been described elsewhere [12,13]. In summary, the 532 nm output of a 20 Hz frequency-doubled Nd:YAG laser is used to pump a dye laser containing a solution of rhodamine 590 laser dye, also known as rhodamine 6G. The dye laser output is frequency-doubled to give UV output, which is directed into an optical fiber and transmitted to a sample where it excites fluorescence. Fluorescence is collected by an optical fiber and transmitted to a monochromator for dispersion. Optical signals are detected by a PMT, whose output is displayed on and analyzed by a digital storage oscilloscope. Instrument control, data acquisition, and data logging are all performed automatically under computer control. Optically-unique laser probes of our own design were fabricated in house. These probes can be embedded in a custom-designed soil-column apparatus (also fabricated in house) or inserted into flasks in the same fashion as conventional laser probes. We typically use five or six probes multiplexed to our optical apparatus.

Several modifications to the apparatus have been performed. Dichroic filters were used to separate the fundamental, frequency-doubled, and frequency-tripled output of the Nd:YAG laser instead of the Pellin-Broca prism used previously. Furthermore, the manually controlled dye-laser doubling crystal was replaced with an automatic crystal (Inrad Autotracker II) in an attempt to increase long-term power stability. The dye cell laser was tuned to 572 nm and its output frequency-doubled to give 286 nm UV light. All fluorescence data were obtained with excitation at this wavelength.

The motor driving the translation stage used to position optical fibers was replaced with a newer model (Newport 850B) along with the new model motion controller (Newport PMC200-P) required to drive it. These changes were made to provide longer travel, more accurate positioning, and hopefully higher reliability.

Probe construction was modified (see Figure 1) to eliminate some problems. The plastic-clad silica fiber formerly used was replaced with silica-clad silica fiber (Fiberguide) to eliminate core pistonning and minimize attenuation of the UV excitation light. Compression fittings had been used to hold the fiber in the body of the probe and to permit rapid replacement of the

fiber should it be damaged. However, those fittings were eliminated from the design because they did not immobilize the fiber as much as was desired and in some cases damaged the fiber cladding and buffer. Consequently, epoxy was used in place of compression fittings. The buffer and cladding around the very tip of the fiber was removed by dissolution in concentrated H_2SO_4 to eliminate any optical heterogeneities produced when the fiber was cut. This also eliminates any void space between the fiber and cladding that could potentially trap fluorescent substances during use which, if present, could result in undesired fluorescence signals. Heat-shrink tubing was placed around the upper part of the probe body and the fiber to minimize fiber flexing at the point where the steel body contacts the fiber—this minimizes the possibility of damaging the fiber cladding. In view of these modifications, the sapphire window covering the end of the fiber was judged to be unnecessary and so was eliminated.

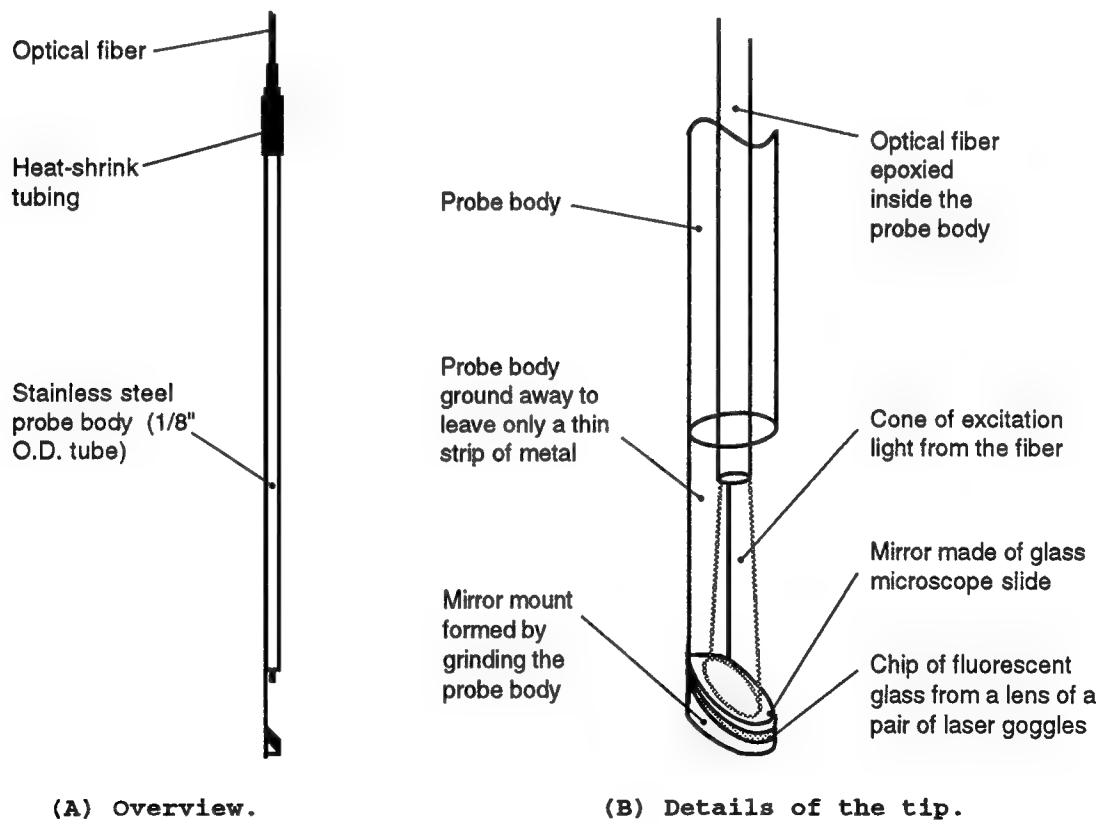


Figure 1. Construction of the laser probe.

Results

Current studies involve the analysis of naphthalene, which is an important component of Air Force jet fuels such as JP-4 and JP-8. Previous experience indicated that photolysis of aqueous naphthalene is a problem with our laser set on higher powers. Naphthalene fluorescence decay from 100 mL of an approximately 100 ppb solution was monitored at 335 nm on an oscilloscope for a short period of time. The drop in intensity suggested that the laser light was photolyzing the naphthalene to a small extent. When the solution was stirred, the intensity rose to its original value. This suggested that fresh naphthalene was being moved into the optical path of the probe. The power was lowered by adjusting the Nd:YAG flashlamp voltage a little. Laser light was then permitted to hit 100 mL of a naphthalene solution (approximately 100 ppb) for approximately six hours. A fluorescence measurement was taken once every six minutes during this time. During this time the 532 nm light from the Nd:YAG laser was found to be stable (as indicated by photodiode measurements), as was the fluorescence from the naphthalene. Figure 2 is a graph of uncorrected naphthalene fluorescence as a function of time. The fact that the fluorescence intensity remained constant

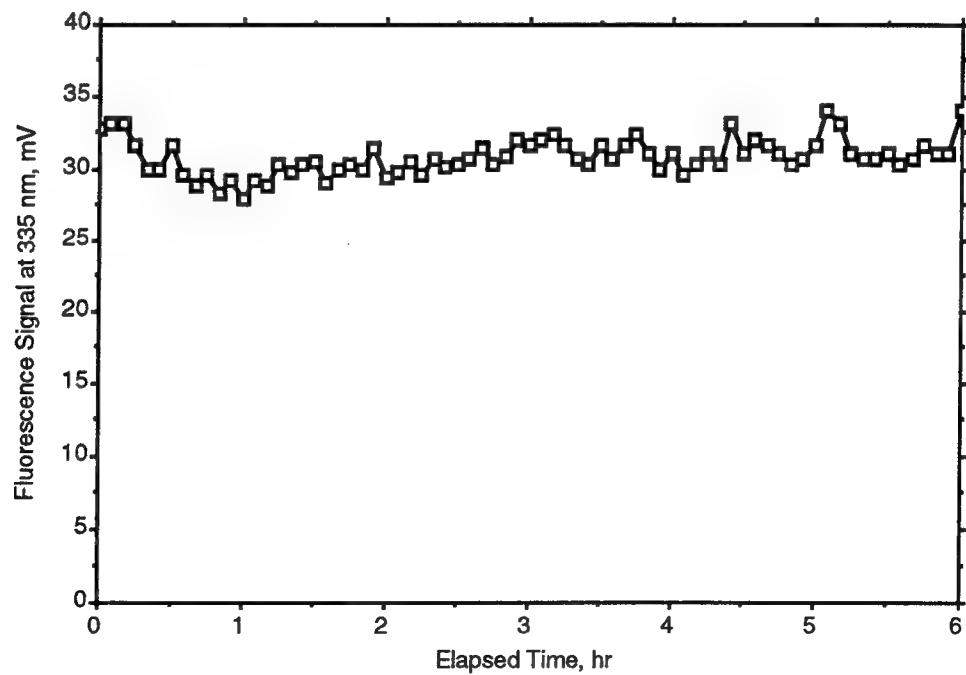


Figure 2. Long-term stability of naphthalene fluorescence.

during the course of the experiment could have meant that the naphthalene was not photolyzed. However, it is also possible that the six-minute sampling interval was long compared to the rate at which fresh naphthalene diffused into the optical path of the probe. Further experiments need to be done to clarify conditions under which photolysis occurs.

Comparison of current data to data of previous studies suggests that the automatic doubling crystal has been of some benefit in stabilizing long-term UV power output. However, this has not solved all of our laser power stability problems, as is evidenced by short-term noise in our fluorescence signals. We tried to power normalize our fluorescence intensities with signals from a photodiode monitoring a fraction of the UV light from the doubling crystal, but this was unsuccessful in eliminating the noise. It is suspected that the active area of the photodiode was too small to give a representative measure of overall beam power (which is probably spatially heterogeneous). Furthermore, the UV light appeared to damage the photodiode. Replacement of the photodiode with a PMT having a much larger active area made little difference other than being much more robust to UV light. Attempts at using the fiber Raman OH signal to eliminate noise were also unsuccessful, as were attempts to use the Raman and PMT signals in tandem as a correction. We also continue to have a less significant but still bothersome problem with long-term drift. Although we found that using the reference signals to compensate for drift was successful sometimes, reproducibility was poor. We found we could satisfactorily eliminate drift by monitoring a reference solution with one laser probe and using the fluorescence intensity from that solution to normalize the signals from the other probes.

A post-doctoral student working in the Fate and Transport Group of the Environics Directorate has developed a mathematical model to describe the interaction between contaminants and layered soil materials in soil columns. We are currently trying to obtain the data necessary to test this postulated model by using multiple laser probes embedded in our soil-column apparatus.

Our first effort in obtaining these data will be with sand that has been screened to provide particle-size homogeneity. Unlike many soil materials, this sand does not contain small particulates that can become suspended and cause light scattering that can interfere with fluorescence measurements.

This will permit relatively quick and easy measurements to be made to get a good idea of the validity of the model.

At least two layers of sand will be put into the soil column. The first layer will be the sand just described. The other layer will consist of a sample of that same sand that has been silanized by the Fate and Transport Group to provide high-carbon surfaces on the particles. It is expected that aqueous naphthalene will be efficiently adsorbed onto those surfaces but that it will not adsorb well onto the untreated sand. Therefore, the untreated and silanized sand will behave as two different kinds of soil materials.

Preliminary studies of aqueous naphthalene equilibrated with treated sand for approximately 24 hours indicated that no adsorption occurred at all. Several silanizing procedures have been attempted with identical results. It is suspected that either the silanizing reagent is bad or that the silanizing procedure being used is ineffective.

If the initial studies with the sand suggest that the model is good, then experiments will be performed with a more complex soil matrix. It is known that the material to be used contains significant quantities of small particles. Our group has previously developed a laser probe that permits light scattering by turbid solutions to be quantified and accounted for when doing fluorescence analysis. That probe will be used to correct for light scattering caused by the mobilization of fine particles from the soil matrix in this study. We intend to finish this series of experiments before proceeding with other research.

Conclusion

Research in the Environics Sensors Program is focused on the development of optical sensors to meet Air Force environmental monitoring needs. Optical fibers coupled to a laser fluorescence spectrometer provide a powerful means of *in situ* analysis.

Accomplishments to date include the development of a unique laser probe that is self-referencing in regard to light scattering. Use of this probe permits quantitative correction to be made for light scattering due to solids suspended in ground water. A complete sensor system has been developed by interfacing a computer and multiple fiber-optic laser probes to a single laser

system. Hardware control and data acquisition and logging are automated and controlled by custom software developed in the Sensors Program. Interaction between fluorescent tracer dyes and soil materials can be evaluated with laser probes embedded in a unique soil-column apparatus developed by the Sensors Group. Current efforts are underway to provide the data needed to test a mathematical model developed by the Fate and Transport Group.

Several areas have been identified as potential subjects for subsequent research. An internal document containing details and a literature review with 147 references has been filed on base [14].

References

1. W. Bains, "Sensors for a Clean Environment," *Bio/Technology*, 10, 515-519 (1992).
2. M.A. Arnold, "Fiber-Optic Chemical Sensors," *Analytical Chemistry*, 21, 1015A-1025A (1992).
3. D.R. Walt, "Fiber-Optic Sensors for Continuous Clinical Monitoring," *Proceedings of the IEEE*, 80, 903-911 (1992).
4. R. Narayanaswamy, "Chemical Transducers Based on Fibre Optics for Environmental Monitoring," *The Science of the Total Environment*, 135, 103-113 (1993).
5. R. Narayanaswamy, "Optical Chemical Sensors: Transduction and Signal Processing," *Analyst*, 118, 317-322 (1993).
6. D.R. Walt, V. Agayn, K. Bronk, and S. Barnard, "Fluorescent Optical Sensors," *Applied Biochemistry and Biotechnology*, 41, 129-138 (1993).
7. G. Frishman and G. Gabor, "Surface Characteristics of Optical Chemical Sensors," *Sensors and Actuators B*, 17, 227-232 (1994).
8. R.E. Dessey, "Waveguides as Chemical Sensors," *Analytical Chemistry*, 61, 1079A-1094A (1989).
9. F.V. Bright, "Modern Molecular Fluorescence Spectroscopy," *Applied Spectroscopy*, 49, 14A-19A (1995).
10. M.D. Barnes, W.B. Whitten, and J.M. Ramsey, "Detecting Single Molecules in Liquids," *Analytical Chemistry*, 67, 418A-423A (1995).
11. I.M. Warner, G. Patonay, M.P. Thomas, "Multidimensional Luminescence Measurements," *Analytical Chemistry*, 57, 463A-482A (1985).
12. B.S. Vogt, T. Tipton, and B.J. Nielsen, "A Multiplexed Fiber-Optic Laser Fluorescence Spectrometer for Fate and Transport Studies of Groundwater Contaminants," *USAF Summer Research Program--1994, Summer Faculty Research Program (SFRP) Reports*, 44-1 to 44-20.
13. B.S. Vogt, "Soil Column Studies With A Fiber-Optic Laser Spectrometer," *USAF Summer Research Program--1993, Summer Research Program Final Reports*, 2 (Armstrong Laboratory), 2-24 to 2-44.
14. The report filed at Tyndall Air Force Base was also entitled "Optical Sensor Research in the Environics Sensors Program." Requests for copies should be submitted to Bruce J. Nielsen, Environics Directorate, Armstrong Laboratory (AL/EQ), Tyndall AFB, FL 32403-6001.

A Recurrent Neural Network for Aircraft Fault Classification

Based on the Bayesian Decision Theory

Jun Wang

Department of Industrial Technology
University of North Dakota
Grand Forks, ND 58202-7118

Final Report for:
Summer Faculty Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC
and
Armstrong Laboratory

August 1995

A Recurrent Neural Network for Aircraft Fault Classification

Based on the Bayesian Decision Theory

Jun Wang

Department of Industrial Technology
University of North Dakota
Grand Forks, ND 58202-7118

Abstract

A Bayesian/neural-network approach is developed for aircraft fault classification. From historic maintenance data, the posterior probabilities of fault classification based on given fault indicator are estimated and derived using the Bayes' rule. Based on the Bayesian decision theory, the fault classification problem is formulated as a linear integer programming problem to minimize an expected loss function with the posterior probabilities. The linear integer programming problem is then convert equivalently to a standard linear programming problem. A two-layer recurrent neural network is used to carry out the computation task for fault classification by solving the formulated linear programming problem. The simulation results of a pilot study based on the historic data on the radar system in F-16 aircraft show that the neural network approach is capable of real-time aircraft fault classification.

A Recurrent Neural Network for Aircraft Fault Classification Based on the Bayesian Decision Theory

Jun Wang

Introduction

An aircraft is a complex electromechanical system composed of thousands of parts.

Because of the complexity aircraft fault diagnosis and classification are challenging tasks, especially for military aircraft for which the cost and time. The traditional diagnostic methods for aircraft maintenance using technical manuals are costly to author and often fail to isolate the cause of the aircraft failure; thus impacting mission readiness and increasing maintenance costs. High field maintenance hours and false removals are often caused by incorrect diagnosis. In addition, historical information from Maintenance Data Collection systems is difficult to access and rarely used. All these factors contribute to the need of a diagnostic system that is capable of learning from historical data to identify the faulty unit and correctly predict false removals, i.e. Bench Check Serviceable or BCS's. Artificial neural networks provide a possible solution to the fault classification problem.

Resembling biological nerve systems more or less in structure, neural networks are parallel distributed models composed of many simple processing elements. In processing information, these elements operate concurrently and collectively. During the past ten year, neural networks for pattern classifications have been one of the most active areas in intelligent systems research. The results of numerous studies have shown the superior performance of neural networks for pattern classification.

Background Information

Bench Check Serviceable (BCS) occurs when the reported faulty component later checked good. BCS occurs at the maintenance shop level and always will result from unnecessary removal of Line Replaceable Units (LRU). Knowledge of whether a BCS occur can help the maintenance technician to modify his diagnostic procedures thus reduces unnecessary removal of LRU's. A well trained technician can identify an LRU BCS from maintenance history and thus provides this project a set of sample data.

Problem Formulation

To facilitate the ensuing explanation, let the numbers of part units, fault categories, and fault indicators be respectively denoted as m , n , and p . Let x_{ij} be denoted as the decision variable defined as follows: for $i = 1, 2, \dots, m, j = 1, 2, \dots, n$;

$$x_{ij} = \begin{cases} 1 & \text{part unit } i \text{ belongs to fault category } j \\ 0 & \text{otherwise} \end{cases}$$

From past data, one can estimate the prior probability of each part unit (e.g., LRU, SRU) belonging to each category and the conditional probability of each fault indicator (e.g., MFL) for any part unit belonging to any category. From past data or the prior and conditional probabilities using the Bayes rule, we can obtain the posterior probability of each part unit belonging to each category given any fault indicator. The prior probabilities

form an m by n matrix. The conditional and posterior probabilities form two m by n by p three-dimensional data arrays.

For each misclassification, there is always some associated cost penalizing the mistake.

The cost can be in terms of dollar amount, time wasted, or combination of these factors.

These cost coefficients form an m by n cost coefficient matrix. Given a fault indicator, the sum of the Schur product of the loss coefficient matrix and the posterior probability matrix corresponding to the fault indicator constitutes the expected cost (loss) function.

By minimizing the expected loss function subject to some feasibility constraints, faults can be classified.

One fundamental feasibility constraint that ensures an LRU can be assigned to only one

class is as follows: $\sum_{j=1}^n x_{ij} = 1, i = 1, 2, \dots, m.$

Given limited resources on expenditure and manpower, another feasibility constraint can

be added: $\sum_{i=1}^m x_{ij} \leq \mu_j$, for $i=1,2,\dots,n$.

The last fundamental constraint is the integrality constraint defining the binary nature of the decision variables: $x_{ij} \in \{0, 1\}$.

In summary, the fault classification problem can be formulated as the following linear programming problem:

For given MFL_k , $k = 1, 2, \dots, p$;

$$\begin{aligned}
\text{maximize} \quad & \sum_{i=1}^m \sum_{j=1}^n P_{ijk} l_{ijk} x_{ij}; \\
\text{subject to} \quad & \sum_{j=1}^n x_{ij} = 1, \quad i = 1, 2, \dots, m; \\
& \sum_{i=1}^m x_{ij} \leq \mu_j, \quad j = 1, 2, \dots, n; \\
& x \in \{0, 1\};
\end{aligned}$$

where $l_{ijk} = w_1 C_{ijk} + w_2 T_{ijk}$.

The above inequality constraint can be easily converted to an equality constraint by adding a slack variable y_j in for each j . Because the coefficients of the constraints have the total unimodular property, the integrality constraint can be replaced with the nonnegativity constraint. The above linear integer programming problem can thus reformulated as the following linear programming problem: Given MFL_k , $k = 1, 2, \dots, p$;

$$\begin{aligned}
\text{maximize} \quad & \sum_{i=1}^m \sum_{j=1}^n P_{ijk} l_{ijk} x_{ij}; \\
\text{subject to} \quad & \sum_{j=1}^n x_{ij} = 1, \quad i = 1, 2, \dots, m; \\
& \sum_{i=1}^m x_{ij} + \mu_j y_j = \mu_j, \quad j = 1, 2, \dots, n; \\
& x_{ij} \geq 0, \quad y_j \geq 0, \quad i = 1, 2, \dots, m; \quad j = 1, 2, \dots, n.
\end{aligned}$$

Dynamical Equation

To Solve the optimization problem, a two-layer recurrent neural network is developed tailored from the deterministic annealing neural network for convex programming (Wang, 1994). The architecture of the two-layer neural network is shown in Figure 1, where the state variables of output neurons and hidden neurons correspond to the decision variables x_{ij} slack variables y_j respectively. For simplicity, the same symbols are used to denote decision/slack variables and corresponding state variables. The dynamical equation of the recurrent neural network is shown as follows:

$$\frac{du_{ij}}{dt} = -\sum_{j=1}^n x_{ij} - \sum_{i=1}^m x_{ij} - \mu_j y_j + 1 + \mu_j - P_{ijk} l_{ijk} e^{-\eta t},$$

$$\frac{dv_j}{dt} = -\mu_j \sum_{i=1}^m x_{ij} - \mu_j^2 y_j + \mu_j^2$$

The standard unipolar sigmoid activation function is used:

$$x_{ij} = g(u_{ij}) = \frac{1}{1 + e^{-\xi u_{ij}}}, \quad y_j = g(v_j) = \frac{1}{1 + e^{-\xi v_j}};$$

where ξ is a scaling constant determining the sensitivity of activation.

Simulation Results

In this specific pilot study, the task is to classify the line replaceable units (LRUs) into three categories: replaceable LRUs, repairable LRUs, and BCS LRUs. The proposed neural network approach, however, can be used for more general classification tasks with any numbers patterns, categories, and evidences.

An MATLAB program has been developed for simulating the recurrent neural network. The testing data is being extracted and processed to estimate the posterior probabilities and cost coefficients.

Concluding Remarks

A Bayesian classifier based on a recurrent neural network is developed for aircraft fault classification. The Bayesian classifier minimizes the expected loss taken both maintenance cost and time into consideration. The proposed recurrent neural network provides a parallel computational model to carry out the optimization task. The proposed Bayesian classifier can serve as a core in a maintenance decision support system for aircraft diagnosis.

References

D. F. Specht, "Probabilistic neural networks," *Neural Networks*, vol. 3, pp. 109-118, 1990.

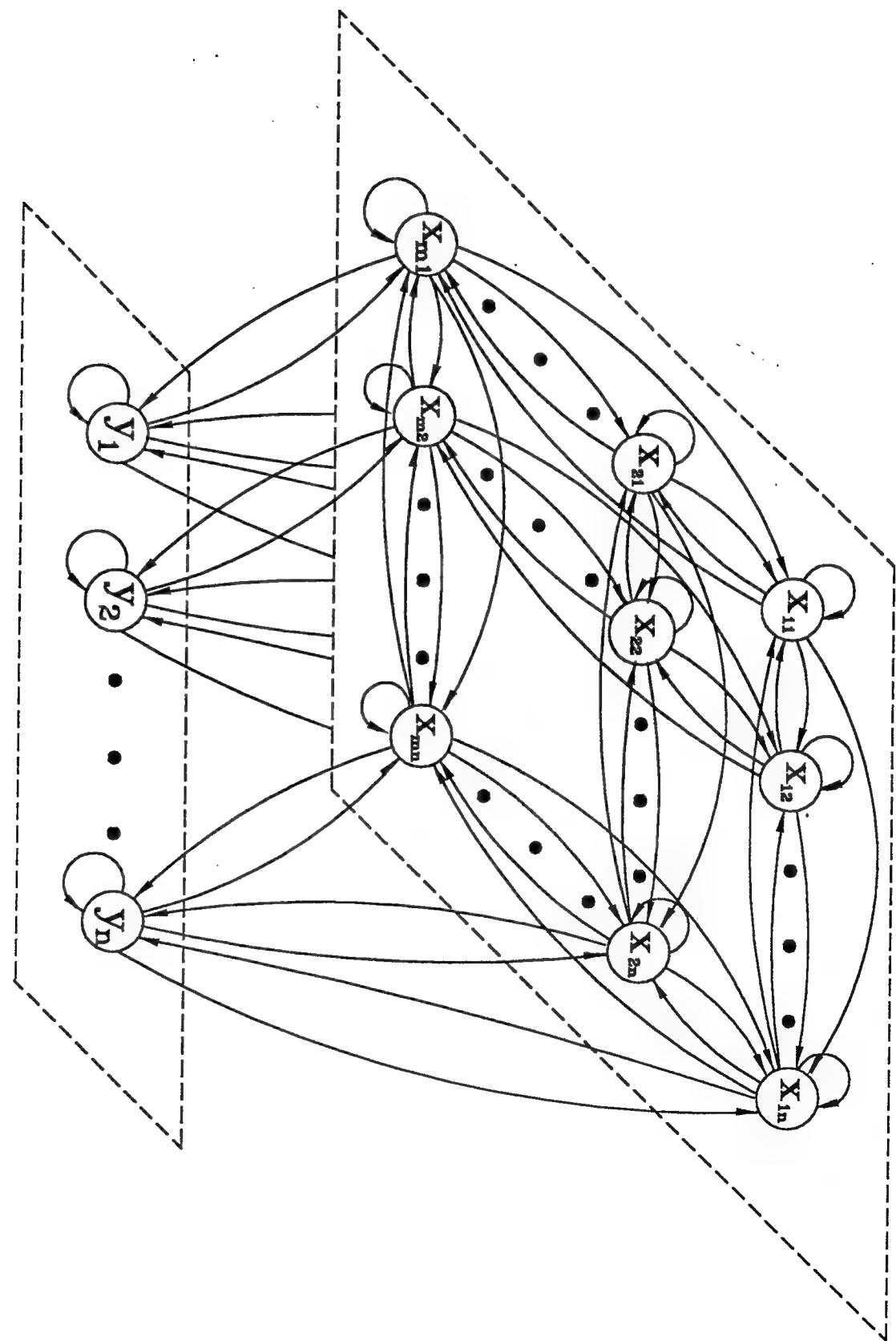
D. F. Specht, "Probabilistic neural networks and the polynomial Adaline as complementary techniques for classification," *IEEE Transactions on Neural Networks*, vol. 1, no. 1, pp. 111-121, 1990.

J. Wang, "Analog neural network for solving the assignment problem," *Electronics Letters*, vol. 28, no. 11, pp. 1047-1050, 1992.

J. Wang, "Analysis and design of a recurrent neural network for linear programming," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 40, no. 9, pp. 613-618, 1993.

J. Wang, "A deterministic annealing neural network for convex programming," *Neural Networks*, vol. 7, no. 4, pp. 629-641, 1994.

J. Wang, "Analysis and design of an analog sorting network," *IEEE Transactions on Neural Networks*, vol. 6, no. 4, pp. 962-971, 1995.



**THERMODYNAMIC EVALUATION OF
EFFLUENT TAILING EFFECTS IN
DISSOLUTION OF BINARY OIL MIXTURES**

**William R. Wise
Associate Professor
Department of Environmental Engineering Sciences**

**University of Florida
A.P. Black Hall
Gainesville, FL 32611-6450**

**Final Report for:
Summer Faculty Research Program
Armstrong Laboratory**

**Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC**

and

Armstrong Laboratory

August 1995

**THERMODYNAMIC EVALUATION OF
EFFLUENT TAILING EFFECTS IN
DISSOLUTION OF BINARY OIL MIXTURES**

William R. Wise
Associate Professor
Department of Environmental Engineering Sciences
University of Florida

Abstract

Local equilibrium with linear interphase partitioning models have failed to describe adequately the dissolution of binary and multi-component oils in groundwater systems. Specifically, the major features not described by such models are persistent tailings of the more soluble components in the effluent concentration histories. To date, the environmental research community has addressed these shortcomings by invoking nonequilibrium models based upon linear partitioning at equilibrium. However, local equilibrium models with nonlinear partitioning have also shown some promise for mimicking tailing phenomena; the degrees of nonlinearity involved can be quite subtle. Critical evaluation of the thermodynamics of nonlinear partitioning in the context of chromatographic theory yields a criterion delineating for binary oils where nonlinear-partitioning induced tailing may be ruled out. Specifically, this condition is met when both components of an oil have enhanced solubility, relative to Raoult's law, as a result of their complementary components. This criterion serves as a proof of the link between nonequilibrium effects and observed tailing phenomena for this class of binary mixtures. Furthermore, such mixtures should be the rule and not the exception, given the nature of partitioning phenomena observed for most mixtures of chemically dissimilar compounds.

THERMODYNAMIC EVALUATION OF EFFLUENT TAILING EFFECTS IN DISSOLUTION OF BINARY OIL MIXTURES

William R. Wise

Introduction

When models describing physical phenomena fail to agree with experimental investigations, those models must be either dismissed or modified to a point where their explanation of observed behavior is brought to an acceptable level. Over the past decade it has become increasingly clear that local equilibrium with linear interphase partitioning models can not explain adequately the leaching of the components from residual saturations of multi-component oils by flowing water. Specifically, the failures are manifest, in part, as persistent tailing of the more soluble components at concentrations from two to four (or more) orders of magnitude below that at which they are observed at early times in the dissolution process. There are two possible modes of failure for the class of local equilibrium with linear interphase partitioning models: failure to establish local equilibrium conditions along streamlines and nonlinear equilibrium interphase partitioning. The observed tailing phenomena could be due to one or the other of these two, or a combination of them.

The environmental research community has taken the position that the failures of these models are due to the lack of achieving local interphase equilibrium along the water flow paths. The various models which have evolved from this position share a common bond: a nonequilibrium process description involving linear interphase partitioning. Among the efforts along these lines are work published by *Borden and Kao* [1992] and *Malone et al.* [1993]. These results are supported anecdotally by dissolution experiments involving residual saturations of single component oils which also exhibit the tailing phenomena (see *Powers et al.* [1992], *Imhoff et al.* [1993], and *Powers et al.* [1994]).

To this author's knowledge, only one group has investigated the possibility that nonlinear partitioning is responsible, at least partly, for the failure of the original model; *Wise et al.* [1992] showed that a local equilibrium model with nonlinear interphase partitioning could be used to mimic the leaching of aromatic hydrocarbons from a commercially available gasoline (data collected by *Borden and Kao* [1989]). The nonlinearities used to describe the interphase partitioning of the aromatic compounds between the oil and water phases were fairly subtle and decreased with decreasing solubility of the compounds. The fitted partitioning relationships became linear as the oil phase concentrations of these aromatics neared zero. By contrast, *Borden and Kao* [1992] used a nonequilibrium (two-site) with linear partitioning model to describe the data with essentially the same results (see Figures 7(a) and 8(a) from the first reference and Figures 5 (a) and (b) from the

second). This is, of course, one of the problems involved with model building: models representing different physical scenarios can be fit to yield "acceptable" results from a minimization of residual errors perspective. Often, examining the underlying phenomena from a different tack may be used to shed insight into the validity of competing models.

This paper delineates where the thermodynamics of oil/water systems might allow nonlinear partitioning into water for the components of binary oil mixtures to produce the tailing effects that are observed for the more soluble component when porous columns containing a residual saturation of such oil are flushed with water. More importantly, a criterion will be offered for systems in which the nonlinear model can not explain the tailing phenomenon for the more soluble component. This condition will be offered as proof that tailing in such systems must be the result of nonequilibrium processes. Such proof has been lacking in the literature to date, despite the proliferation of nonequilibrium models. The possibility of powering up this criterion to multi-component oils is then discussed.

On the Prevalence of the Linear Equilibrium Partitioning Model

Efforts to examine the environmental impact of multi-component oils such as petroleum inadvertently introduced into the subsurface have relied on the premise that equilibrium partitioning of a compound, say benzene, for example, between oil and water phases is linear over the entire domain of possible mole fractions. Thermodynamically, this is known as Raoult's law. *Abriola and Pinder* [1985], *Corapcioglu and Baehr* [1987], and *Mercer and Cohen* [1990], to name a few, reviewed this approach. *Cline et al.* [1991] demonstrated that Raoult's law could be used to describe adequately the aqueous phase concentrations of aromatic constituents of unweathered gasolines in contact with water. They did not, however, investigate the validity of Raoult's law as the compositions of the gasolines changed, as occurs in the dissolution process. It is as the composition changes where partitioning nonlinearities could induce tailing phenomena. Partitioning models such as UNIQUAC and UNIFAC, based upon chemical interactions, are available to predict interphase distributions of chemical species over all possible compositions [Walas, 1985]. These models may be used to supplement experimental partitioning studies for mixtures of interest. *Burris and MacIntyre* [1986] describe appropriate protocols for such experiments.

Data presented by *Fried et al.* [1979] and *Borden and Kao* [1992] indicate that modeling the partitioning of certain aromatic components (benzene, toluene, ethylbenzene, and the three isomers of xylene, a.k.a. BTEX) between oil (gasoline) and water phases with a simple linear equilibrium partitioning expression is not valid. Specifically, tails in the aqueous concentration histories of the more soluble, aromatic constituents of the oil phase during water flushing found in the above

mentioned data are incongruous with the linear, equilibrium model. To address this shortcoming, *Borden and Kao* [1992] included a kinetic effect focusing on a rate limited mass transfer mechanism. They conceptually separated the oil into "fast" and "slow" portions with different mass transfer characteristics and used the ratio of these portions, in part, to fit their data.

Recent investigations into kinetic effects of interphase processes include those by *Miller et al.* [1990], *Powers et al.* [1991], and *Brusseau* [1992]. The first team examined the interphase mass transfer of residual toluene into flushing water within a column of packed glass beads. Mass transfer rates were found to be correlated to the aqueous phase velocity, but not to the size of the glass beads. Equilibrium was shown to be achieved rapidly over a variety of residual toluene saturations and aqueous phase velocities. The second team presented arguments that the amount of interfacial area exposed to flowing water is a rate-limiting factor in the dissolution of an immobile, single-component, immiscible oil. Basically, their work addressed the problem of physical access to the residual blobs. The third work examined the kinetic effects that may be attributed to four cases: (1) immiscible liquid residing in the nonadvective domain, (2) immiscible liquid residing in the advective domain, (3) two-resistance liquid-liquid mass transfer (diffusion within the immiscible fluid and interfacial mass transfer limitations), and (4) single-resistance liquid-liquid mass transfer (interfacial mass transfer limitations).

Background

The present paper steps back to where the choice was made to pursue kinetic models of dissolution for binary and multi-component oils and to critically evaluate this decision. One principle objective of this paper is to identify systems in which this decision is soundly supported by the nature of the equilibrium partitioning, even when it is nonideal (nonlinear). A brief review of how partitioning nonlinearities might affect the dissolution of binary and multi-component oils follows.

Chromatographic Interpretation of Partitioning

Wise et al. [1990 and 1992] presented a chromatographic interpretation of the partitioning of a compound between residual oil and mobile water phases. Their work did not address the issue of equilibrium, but instead focused on the distribution of the compound, its concentrations in both phases, without regard to the system's free energy. They concluded that effluent concentration histories of aromatic compounds leaching from gasoline could be simulated effectively using a local equilibrium model based upon a favorably nonlinear partitioning relationship between the residual gasoline and a flushing water phases, as that illustrated in Figure 1. Their conclusion is supported by quantitative analysis of data presented by *Borden and Kao*

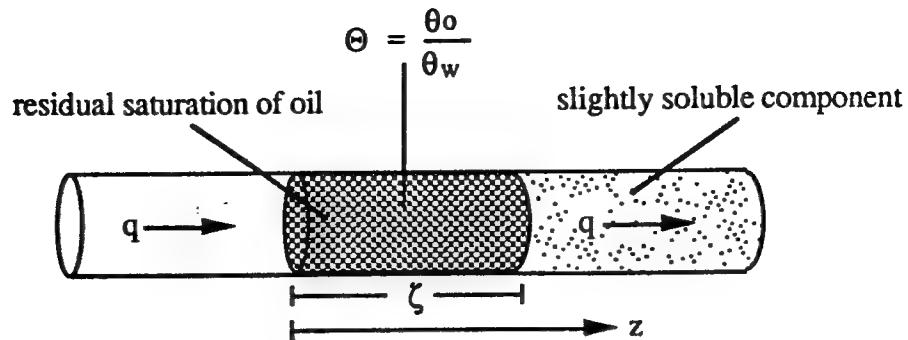


Figure 1. Schematic of water flushing through a layer of residual oil of thickness ζ .

[1992] (who, remember fit the same data to a two-site kinetic model) and qualitative analysis of data presented by *Fried et al.* [1979]. An aqueous effluent (at the downstream extent of the residual fuel zone) concentration history of toluene, from *Borden and Kao* [1992], is shown in Figure 2(a), illustrating the tailing effect. *Wise et al.* [1992] modeled this data (shown as the solid

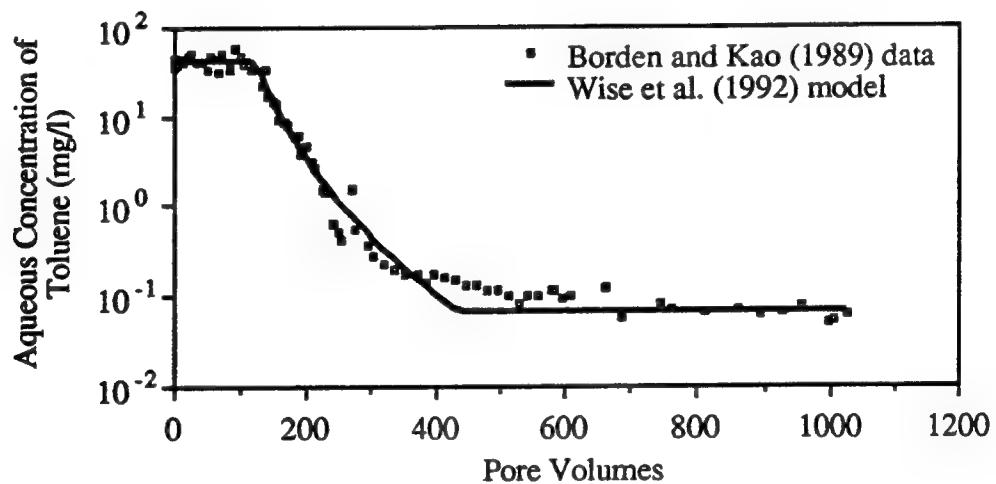


Figure 2(a). Local equilibrium with nonlinear partitioning model fit to toluene data from *Borden and Kao* [1989] experiment: column data. (After *Wise et al.* [1992].)

line in Figure 2(a)) using a method of characteristics solution of the following advective (It is appropriate to ignore dispersion in such systems as the number of pore volumes over which the transition from the higher to the lower effluent concentrations occurs is quite large.) streamline transport equation:

$$\zeta \left[1 + \Theta \frac{\partial c_o}{\partial c_w} \right] \frac{\partial c_w}{\partial \tau} + q \frac{\partial c_w}{\partial z} = 0, \quad (1)$$

where c_w [ML⁻³] and c_o [ML⁻³] are the aqueous and oil phase concentrations of the compound, respectively, Θ is the ratio of the oil content (θ_o) to water content (θ_w) in the region contaminated by the residual fuel, and τ is the number of pore volumes of water flushed past the oil of axial extent ζ [L], at a Darcy flux of q [LT⁻¹] along the axis of the stream tube, z [L]. The nonlinearity they fit is a partitioning relation concave toward the c_w axis in the c_w - c_o plane (see Figure 2(b)). Only near the origin of the c_w - c_o plane (between three and four orders of magnitude below the maximum observed concentration for benzene), does the concavity decay. This nonlinearity is responsible for the curved portion of the model fit in Figure 2(a). The second, lower plateau of the model fit is a manifestation of the partitioning becoming linear at (and below) the plateau concentration. Similar results hold for the other BTEX compounds. The deviations from linearity are positively correlated to the pure component solubilities of the BTEX compounds, suggesting that there may exist a thermodynamic basis for these observed partitioning phenomena.

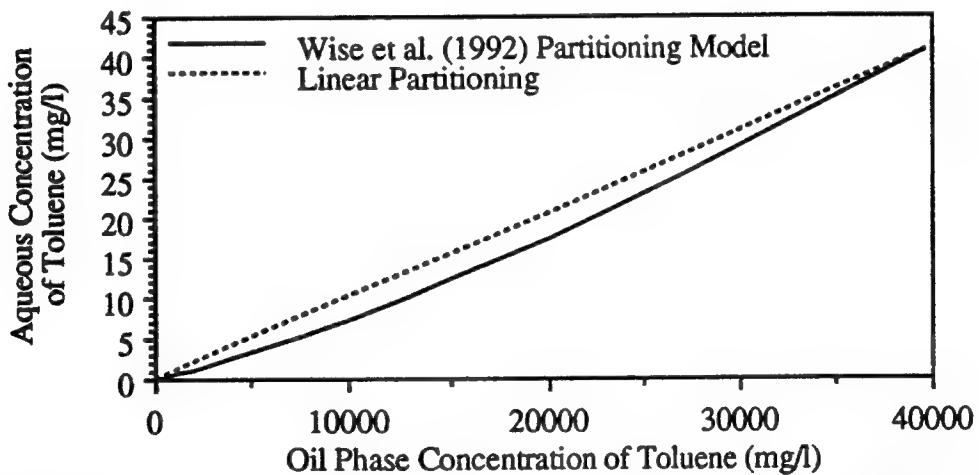


Figure 2(b). Local equilibrium with nonlinear partitioning model fit to toluene data from *Borden and Kao* [1989] experiment: inferred partitioning. (After *Wise et al.* [1992].)

From the modeling perspective, the chromatographic nature of water moving past the residual oil, combined with only a slightly favorable partitioning nonlinearity (which becomes linear at the origin), results in persistent effluent tailing, even in the absence of kinetic limitations. This hypersensitivity of the modeled effluent response relative to the partitioning makes the issue of nonlinearity important. Note that tailing is only anticipated for favorable interphase partitioning relationships; favorable in this sense means that the component of interest favors the oil phase at

low concentrations. Mathematically, this is manifest as a concavity in the partitioning relationship toward the c_w axis when plotted in the c_w - c_o plane. For components whose partitioning functions are either concave toward the c_o axis or linear, the tailing effect is not seen in the solution to equation (1), subject to the appropriate initial and boundary conditions (see Wise *et al.* [1992]).

Theory

This section focuses upon the thermodynamics of equilibrium partitioning between phases.

Concept of Fugacity

Fugacity is a measure of the tendency of a substance to escape by some chemical process from the phase in which it exists into another. The fugacity, f [ML⁻¹T⁻²], of a gas is defined by:

$$\frac{\lambda}{\lambda^\theta} = \frac{f}{P^\theta}, \quad (2)$$

where λ [ML⁻³] and λ^θ [ML⁻³] are the activities of the gas at the actual and standard states, respectively, and P^θ [ML⁻¹T⁻²] is the vapor pressure (pressure of the equilibrated liquid-vapor system) of the gas at the standard state.

c_w as a Measure of Fugacity

The ratio of the absolute activities of species i in any two coexistent liquid phases, α and β , is equal to the ratio of the fugacities of the species in a gas phase in equilibrium with α and β , respectively [Guggenheim, 1967, p. 181]. The activity of species i in phase β may be taken as a measure of the fugacity of species i in phase α at a specified temperature and pressure. (The fugacity, f_i , of species i in a liquid phase is defined by $f_i/\lambda_i =$ a constant at a given temperature for variable pressure, where λ_i is the activity of i in the liquid mixture.) For this work, the aqueous phase (phase β) concentration (which approximates the activity in the aqueous phase, especially at lower concentrations) is used to describe the fugacity of oil constituents in the oil phase (phase α). At this point there is essentially a question as for which phase, if either, Henry's law is applicable. "It is a fair assumption that the activity coefficient (in water) remains constant in the dilute concentration range from zero to saturation, provided that the saturation or solubility limit is less than a mole fraction of 1%" [Mackay *et al.*, 1991]. This result follows from the fact that at such low concentrations, solute-water interactions dominate solute-solute interactions, so there is a constant energy of association for added solute molecules. As a result, application of Henry's law is justified in the aqueous phase. Therefore, the aqueous phase concentration is used as an approximate measure of activity in the aqueous phase (the activity coefficient is considered constant

in the aqueous phase) and consequently the fugacity of components in the fuel phase, from which they "escape." Consequently, the mixture which is examined is the oil phase, not the aqueous phase, which serves as an indicator for the fugacities of the components comprising the oil phase. Of course, from an environmental perspective, the aqueous phase is of major concern.

Duhem-Margules Relation for Binary Mixtures

The Gibbs-Duhem relation (at constant temperature and pressure--conditions assumed for the remainder of this paper) is given by:

$$\sum_i x_i^\alpha d\mu_i^\alpha = 0, \quad (3)$$

where x_i^α is the mole fraction of species i in phase α and μ_i^α [$ML^2T^{-2}mole^{-1}$] is the chemical potential of species i in phase α . Essentially, the Gibbs-Duhem relation documents that although there are $c+2$ intensive quantities describing phase α , T [K], P [$ML^{-1}T^{-2}$], and μ_i^α , where c is the number of components comprising phase α , there are only $c+1$ degrees of freedom associated with the phase (apart from the size of the phase, which is irrelevant to intensive properties). This result follows from the Gibbs phase rule. As written above, the Gibbs-Duhem relation describes the relation between the possible changes in chemical potential of the component species comprising phase α .

By definition, the chemical potential, μ_i , and the absolute activity, λ_i , of species i are related by:

$$\mu_i = RT \ln \lambda_i, \quad (4)$$

where R is the universal gas constant ($R = 8.3143$ J/mole-deg). Using equations (2), (3), and (4), at a constant temperature and pressure, the Gibbs-Duhem relation is equivalently stated:

$$\sum_i x_i^\alpha d\ln \lambda_i^\alpha = 0. \quad (5)$$

Equation (5) is known as the Duhem-Margules relation. In the case of a binary mixture:

$$x_1 \frac{\partial \ln f_1}{\partial x_2} + x_2 \frac{\partial \ln f_2}{\partial x_2} = 0, \quad (6)$$

where f_1 and f_2 are the fugacities of species 1 and 2, respectively in the mixture. Obviously, there exists the constraint:

$$x_1 + x_2 = 1. \quad (7)$$

Note that the derivatives are taken in the same direction, here with respect to x_2 . If one component's fugacity is plotted versus the mole fractions, the other component's behavior is determined by the above equation. It is sometimes more convenient to express the Duhem-Margules formula in the following form, recognizing that $dx_1 = -dx_2$:

$$-x_1 \frac{\partial \ln f_1}{\partial x_1} + x_2 \frac{\partial \ln f_2}{\partial x_2} = 0. \quad (8)$$

Near Zero Behavior - Raoult's Law for Component 1; Henry's Law for Component 2

As component 2 is depleted, $x_1 \rightarrow 1$ while $x_2 \rightarrow 0$. Raoult's law becomes valid for component 1:

$$\frac{f_1}{f_1^0} = x_1. \quad (9)$$

where f_1^0 is the fugacity of pure component 1 in the reference state. Taking the natural logarithm of equation (9) and differentiating the result yields:

$$\frac{\partial \ln f_1}{\partial x_1} = \frac{1}{x_1}. \quad (10)$$

Substituting this expression into the Duhem-Margules relation for a binary mixture, equation (8), yields:

$$x_2 \frac{\partial \ln f_2}{\partial x_2} = 1. \quad (11)$$

Which may be integrated to yield Henry's law for component 2:

$$f_2 = kx_2, \quad (12)$$

where k [$ML^{-1}T^{-2}$] is a constant. A near-zero mole fraction for one component results in linear equilibrium partitioning for both components. When the same methodology used above is applied to slightly higher x_2 (lower x_1) values, subtle information about partitioning nonlinearities may be obtained from the gross partitioning behavior that is more easily observed.

Four Possible Nonlinearity Patterns for Binary Oil Mixtures

On a fundamental level, each component of a nonideal binary mixture may have either an enhanced or a depressed activity in the aqueous phase in comparison to that predicted by Raoult's law. Components that have regions exhibiting both types of behavior are not considered presently. Table 1 summarizes the four possible patterns of nonlinear partitioning for binary oil mixtures. The four cases of interest are: (a) components 1 and 2 both have enhanced solubility, (b) component 1 has enhanced solubility while component 2 has depressed solubility, (c) component 1 has depressed solubility while component 2 has enhanced solubility, and (d) both components have depressed solubilities. Below, the Duhem-Margules equation is used to bring to light some of the more subtle features of nonlinear partitioning for these classes of binary mixtures.

Table 1. Four Cases of Binary Oil Nonidealities

Case:	(a)	(b)	(c)	(d)
Component 1 Solubility compared to Raoult's law	enhanced	enhanced	depressed	depressed
Component 2 Solubility compared to Raoult's law	enhanced	depressed	enhanced	depressed

Applying the Duhem-Margules Approach to Nonideal Binary Mixtures

A binary mixture of a less soluble compound (component 1) and a more soluble compound (component 2) is considered. Therefore, component 2 is preferentially leached from the oil by water in comparison to component 1. Note that the binary mixture does not address any water dissolved in the oil phase. This approximation is not anticipated to affect the results of the present analysis (see *Leinonen and Mackay [1973]*).

Equilibrium partitioning for the four cases over the domains of possible mole fractions are conceptually shown in Figures 3(a)-(d). The left and right ordinates represent the aqueous phase concentrations of the components nondimensionalized to their solubilities in water. The dashed lines correspond to Raoult's law for the components. The near-zero behavior discussed above is

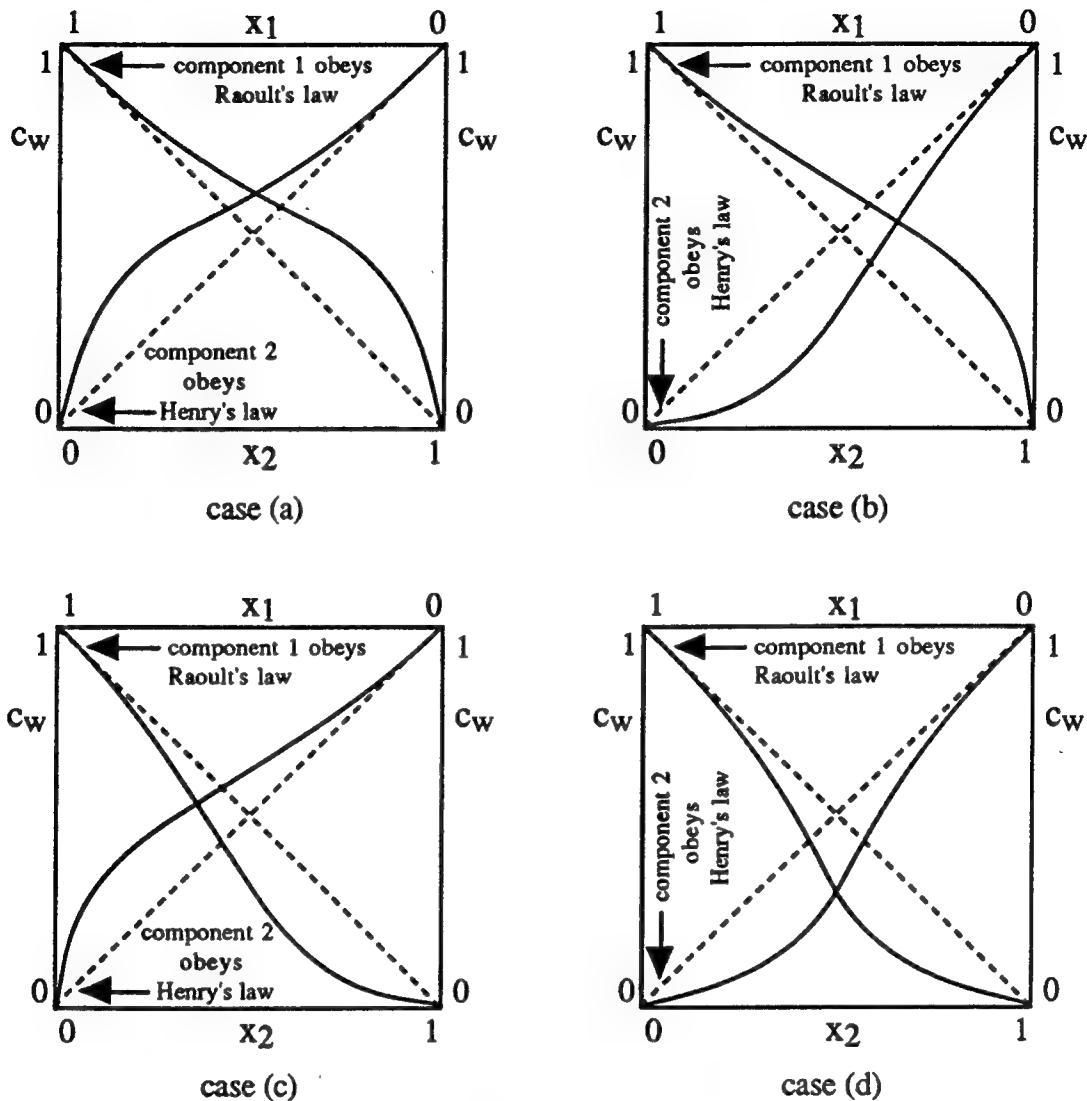


Figure 3. Graphical representation of Duhem-Margules relation for nonideal binary mixtures. (Component 2 is more soluble than component 1.)

valid for each mixture, as indicated in Figures 3(a)-(d). For each case, as the mole fraction of component 2, x_2 , is increased from zero, its partitioning behavior is determined by analyzing the behavior of component 1. For both compounds, the near-zero behavior is linear. However, only for component 1 is the slope known; given by Raoult's law (the slope is equal to unity in the nondimensionalized space). The slope for Henry's law for component 2, the integration constant k in equation (12), is empirically determined as infinite dilution is approached.

Raoult's law for the component 1 (each case) implies that at a mole fraction, x_1 , of unity:

$$\frac{\partial \ln f_1}{\partial x_1} = \frac{1}{x_1}, \quad (13)$$

as stated above in equation (10). As component 1 deviates from Raoult's law, the value of $\frac{\partial \ln f_1}{\partial x_1}$ must either decrease or increase to accommodate the enhanced or depressed solubility of component 1 in that mixture, respectively:

$$\begin{array}{llll} \frac{\partial \ln f_1}{\partial x_1} < \frac{1}{x_1}; & \frac{\partial \ln f_1}{\partial x_1} < \frac{1}{x_1}; & \frac{\partial \ln f_1}{\partial x_1} > \frac{1}{x_1}; & \frac{\partial \ln f_1}{\partial x_1} > \frac{1}{x_1}. \\ \text{case (a)} & \text{case (b)} & \text{case (c)} & \text{case (d)} \end{array} \quad (14)$$

It follows from equation (8) and relation (14) that:

$$\begin{array}{llll} \frac{\partial \ln f_2}{\partial x_2} < \frac{1}{x_2}; & \frac{\partial \ln f_2}{\partial x_2} < \frac{1}{x_2}; & \frac{\partial \ln f_2}{\partial x_2} > \frac{1}{x_2}; & \frac{\partial \ln f_2}{\partial x_2} > \frac{1}{x_2}. \\ \text{case (a)} & \text{case (b)} & \text{case (c)} & \text{case (d)} \end{array} \quad (15)$$

The inequalities are preserved during integration and exponentiation to yield:

$$\begin{array}{llll} f_2 < kx_2; & f_2 < kx_2; & f_2 > kx_2; & f_2 > kx_2, \\ \text{case (a)} & \text{case (b)} & \text{case (c)} & \text{case (d)} \end{array} \quad (16)$$

respectively, where k is the constant for Henry's law given in equation (12). Note that deviations from Raoult's law for component 1 (relation (14)) and Henry's law for component 2 (relation (15)) must begin at the same composition within a given mixture. Each of the four cases warrants a separate discussion.

Case (a)

Case (a) corresponds to enhanced solubility for both components. Figure 4(a) illustrates the near-zero behavior for component 2. Note that since this component's solubility is enhanced in the gross sense, Henry's constant is larger than the slope of Raoult's law. Relation (16) for this case indicates that the partitioning must be concave toward the axis corresponding to the mole fraction of this component, that is, the x_2 axis (this is essentially the same as the c_0 axis discussed above). (This concavity will hold up unto an inflection point required for this partitioning trace to achieve Raoult's law as x_2 nears unity.) Chromatographic theory indicates that this (unfavorable) partitioning relationship will not result in tailing of component 2 in the effluent concentration history. This is of particular importance because chemical dissimilarity generally (but not always)

causes the activity coefficients in the oil phase to exceed unity [Mackay *et al.*, 1991]. Consequently, case (a) is taken to be the rule; the other cases are taken as the exceptions. Systems in which Raoult's law is followed (for both components, as constrained by the Duhem-Margules relation) will behave in the same manner as mixtures corresponding to case (a). Consequently, from the performance standpoint, ideal behavior could be included in case (a).

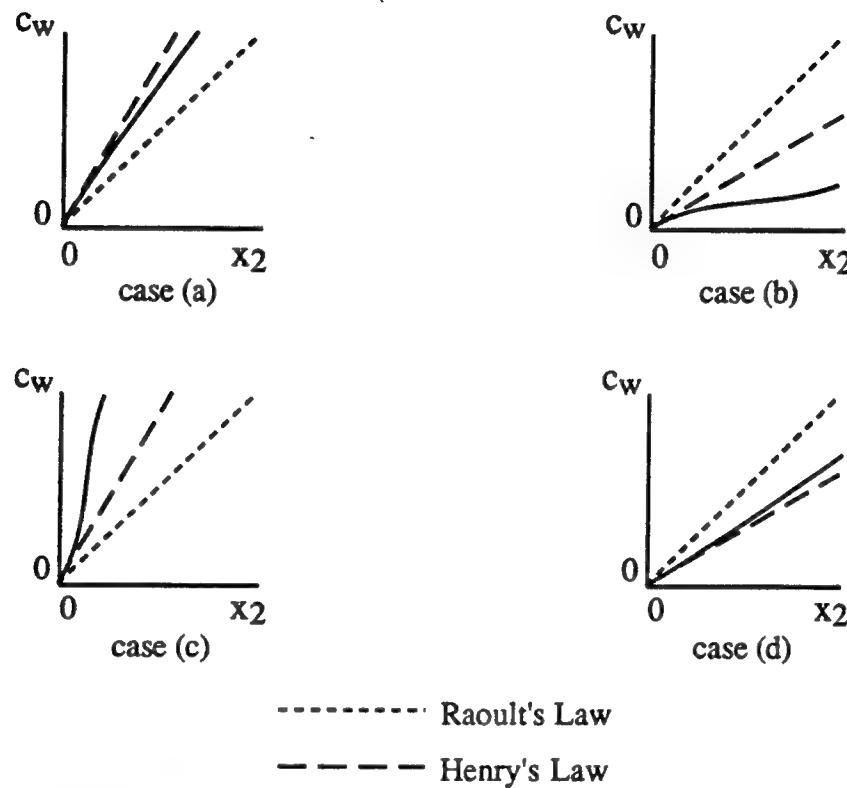


Figure 4. Graphical representation of near zero (component 2) behavior of the Duhem-Margules relation for nonideal binary mixtures. (Component 2 is more soluble than component 1.)

Case (b)

This case differs from the previous one in that the solubility of component 2 is depressed, rather than enhanced. Consequently, for this component, Henry's law will lie below Raoult's law as indicated in Figure 4(b). Relation (16) again implies that as the behavior deviates from Henry's law, the concavity is toward the axis corresponding to the mole fraction for this component. However, in order for the trace of the partitioning of component 2 throughout the x_2 - c_w space to reach the point (1, 1), there must exist an inflection point changing the concavity from the x_2 axis to the c_w axis. Another inflection point is required for this component to approach Raoult's law as x_2 nears unity. Between the two inflection points, the concavity is consistent with potential

nonlinear partitioning induced tailing phenomena (it is favorable). That is, this possibility may not be dismissed summarily, as it may for case (a).

Case (c)

For case (c), the solubility for component 1 is depressed and that for component 2 is enhanced. Figure 4(c) illustrates that, for this case, Henry's law resides above Raoult's law. Relation (16) implies that the near-zero trace of component 2's partitioning is concave toward the c_w axis, consistent with potential nonlinear partitioning induced tailing phenomena. Again this possibility may not be negated for systems exhibiting case (c) type behavior. In fact, this type of concavity melding into the Henry's law region is consistent with the partitioning model used by *Wise et al.* [1992] to mimic the effluent concentrations of aromatic compounds as they were leached from a residual saturation of petroleum. (Again, two inflection points will be required to be in concordance with this concavity and the ultimate need to approach Raoult's law as x_2 nears unity.) This case is of great interest because the enhanced solubility of component 2 appears to produce a nominally unfavorable partitioning relationship for that component. However, when the gross behavior is examined in light of the Duhem-Margules equation, the partitioning of component 2 is readily seen to be favorable at concentrations immediately above the domain where Henry's law is obeyed. It is quite likely that such a subtle effect (see Figure 4(c)) might be missed by experimental data collection regarding the partitioning followed by data regression. The low concentrations at which the pertinent (favorable) nonlinearity might be found would likely be in an analytical region fraught with significant relative error bars. Furthermore, a predetermined form of the partitioning relationship would likely not include this effect; a Freundlich-type fit, for example, would most likely capture the concavity over the unfavorable portion of the domain (between inflection points) and extrapolate this behavior to the origin. It must be kept in mind that such empirically-based models do not have to obey thermodynamics the way the actual systems do.

Case (d)

The solubilities of both components are depressed in case (d). The near zero behavior of component 2 is illustrated in Figure 4(d). As in case (b), Henry's law for component 2 resides below Raoult's law. Relation (16) indicates that the concavity is, like in case (c), toward the c_w axis (it is favorable). Therefore nonlinear partitioning induced tailing phenomena are also possible for systems described by this case. (One inflection point is required to achieve Raoult's law as x_2 nears unity.)

Review of the Four Cases

Only in case (a), where the solubilities of both components are enhanced by their mutual presence, can nonlinear partitioning induced tailing phenomena be dismissed *a priori*. This result has been proven using a straightforward thermodynamic analysis. Consequently for binary oil mixtures characterized by case (a), observed tailing phenomena may be taken as clear evidence of nonequilibrium processes governing dissolution. Mixtures characterized by cases (b), (c), and (d) do not carry with them the same guarantee. For dissolution of such oils to be considered not to be at local equilibrium, the burden of proof lies outside of the realm of equilibrium partitioning thermodynamics.

Using the present conceptual analysis, there is no way to establish where the deviations from Raoult's and Henry's laws begin for binary oil mixtures, only that they exist. The exact locations of the inflection points are also unknown, only that they must exist. Activity models, such as UNIFAC, may shed some insight into the approximate behavior of a given binary-oil/water system. While UNIFAC is only a model, it probably may be used to develop the gross behavior from which to classify a given binary oil mixture into one of the four cases (or alternatively an ideal mixture where both components obey Raoult's law over the entire composition domain).

One Example Binary System: Case (a)

Flushing experiments were performed by *Smith* [1994] in sand columns in which residual saturations of a mixture of decane (component 1) and toluene (component 2) were emplaced. Effluent concentration histories for toluene are presented in Figure 5(a) for four replicate columns. Note the persistent tailings of toluene in the effluent histories. Figure 5(b) illustrates relevant partitioning information for this system. Included in this figure are comparisons of the UNIFAC predictions for partitioning with Raoult's law for both components. The system is clearly predicted to be a case (a) example. Supporting this is a fit for the toluene partitioning also illustrated in the figure, derived from data spanning toluene mole fractions from 3×10^{-6} to 0.3. Note the excellent agreement with the UNIFAC prediction. Unfortunately, decane is not as amenable to analysis as toluene, so no partitioning experiments were attempted for it. Given the arguments developed above, the tailing phenomena presented in Figure 5(a) may be attributed to kinetic effects with a great deal of confidence.

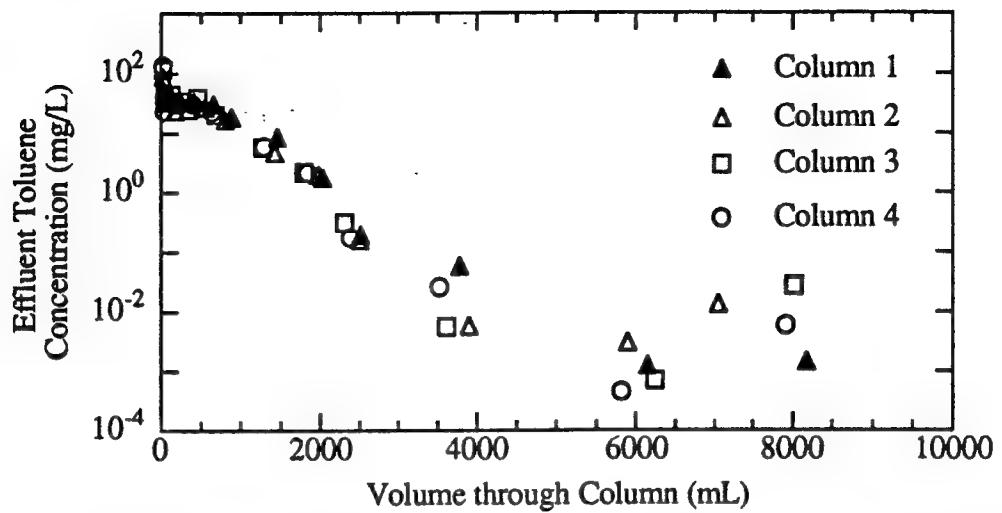


Figure 5(a). Effluent concentration histories from four sand-filled columns, each containing a residual saturation of a mixture of decane (initial mole fraction of 0.924) and toluene (initial mole fraction of 0.076) studied by *Smith* [1994].

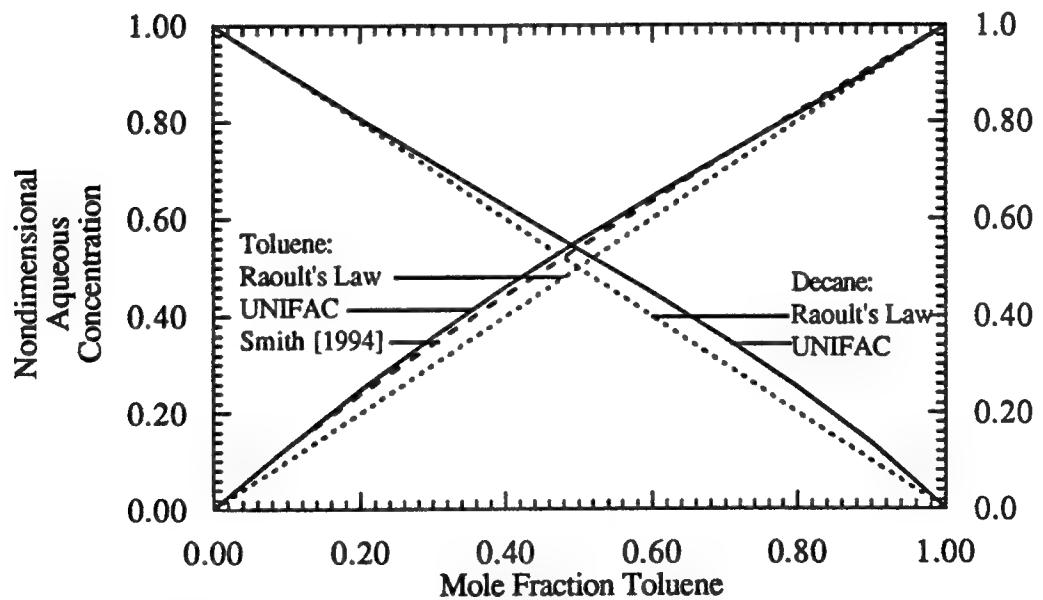


Figure 5(b). Equilibrium partitioning predictions for decane-toluene mixture studied by *Smith* [1994].

On Trying to Extend the Duhem-Margules Approach to More Complex Oil Mixtures

For an oil comprised of n components, there exist $n - 1$ degrees of freedom within the composition space. As a result, as the dissolution process occurs, a change in the mole fraction for a given compound is given by the projection of the composition path vector onto the (mole fraction) axis for that component. By definition, at most one component could be in its Raoult's law region at a time. The union of all of the domains where Raoult's law is valid for the various components is a region extending into the composition space from the boundary of the space. At an arbitrary point along the composition path corresponding to the preferential removal of a particular component from a given multi-component mixture, there is no restriction that Raoult's law must be obeyed for any component unless the path traverses through a part of the aforementioned boundary region. As a result, there is no *a priori* information on the concavities of the remaining components along the composition path vector. (Remember from the binary mixture analysis, that the concavities of these partitioning relations were seen to change over the composition domain in order to accommodate the thermodynamic constraints.) Essentially, simple knowledge of whether the solubilities are enhanced or depressed does not shed any light into the chromatographic behavior under local equilibrium conditions for multi-component (more than two) mixtures. Therefore, the Duhem-Margules equation, equation (5), can not be manipulated simply to yield the qualitative nature of any nonideal partitioning associated with the component of interest for a multi-component oil in the manner in which it can for binary mixtures. Of course, if the partitioning for all components of a mixture are known to a high level of detail, this information could be used to deduce the possibility of nonlinear partitioning induced tailing, for given species, under the constraint of local equilibrium.

Discussion

For binary oil mixtures where both components' solubilities are enhanced by the presence of each other (case (a) above), observed tailing may be attributed with total confidence to nonequilibrium processes. This scenario should be the rule, as chemical dissimilarities tend to produce such equilibrium partitioning behavior. Proof of this fact has been heretofore lacking.

Unfortunately, more complicated multi-component mixtures are not readily analyzed by the same method used to investigate binary mixtures. No criterion is offered to guarantee that nonequilibrium effects are totally responsible for tailing effects for multi-component oils. However, with support for the kinetic limitations of case (a) binary-oil/water systems, it becomes even more compelling to continue to employ kinetic models to processes involving the dissolution of multi-component oils.

Examination of the partitioning rules for binary components brings to light some interesting subtleties not traditionally considered in engineering studies. These subtleties are pertinent to championing the kinetic-type models over the nonlinear equilibrium partitioning model, despite the successes for both in mimicking observed behavior. It must be remembered that model success and accuracy are not one and the same.

References

Abriola, L.M., and G.F. Pinder. 1985. A Multiphase Approach to the Modeling of Porous Media Contamination by Organic Compounds, 2, Numerical Simulation. *Water Resources Research.* v. 21, no. 1, pp. 19-26.

Borden, R.C., and C.M. Kao. 1989. Water Flushing of Trapped Residual Hydrocarbon: Mathematical Model Development and Laboratory Validation. *Proceedings of The Petroleum Hydrocarbons and Organic Chemicals in Ground Water: Prevention, Detection and Restoration.* NWWA. pp. 175-189.

Borden, R.C., and C.M. Kao. 1992. Evaluation of Groundwater Extraction for Remediation of Petroleum-Contaminated Aquifers. *Water Environment Research.* v. 64, no. 1. pp. 28-36.

Brusseau, M.L. 1992. Rate-Limited Mass Transfer and Transport of Organic Solutes in Porous Media That Contain Immobile Immiscible Organic Liquid. *Water Resources Research.* v. 28, no. 1, pp. 33-45.

Burris, D.R., and W.G. MacIntyre. 1986. A Thermodynamic Study of Solutions of Liquid Hydrocarbon Mixtures in Water. *Geochimica et Cosmochimica Acta.* v. 50, pp. 1545-1549.

Cline, P.V., J.J. Delfino, and P.S.C. Rao. 1991. Partitioning of Aromatic Constituents into Water from Gasoline and Other Complex Solvent Mixtures. *Environmental Science and Technology.* v. 25, no. 5, pp. 914-920.

Corapcioglu, M.Y., and A.L. Baehr. 1987. A Compositional Multiphase Model for Groundwater Contamination by Petroleum Products, 1, Theoretical Considerations. *Water Resources Research.* v. 23, no. 1, pp. 191-200.

Fried, J.J., P. Muntzer, and L. Zilliox. 1979. Ground-Water Pollution by Transfer of Oil Hydrocarbons. *Ground Water.* v. 17, no. 6, pp. 586-594.

Guggenheim, E.A.. 1967. *Thermodynamics.* Elsevier Science Publishers, Amsterdam, The Netherlands. 390 pp.

Imhoff, P.T., P.R. Jaffé, G.F. Pinder, An Experimental Study of Complete Dissolution of a Nonaqueous Phase Liquid in Saturated Porous Media. 1993. *Water Resources Research.* v. 30, no. 2, pp. 307-320.

Leinonen, P.J., and D. Mackay. 1973. The Multicomponent Solubility of Hydrocarbons in Water. *Canadian Journal of Chemical Engineering.* v. 51, pp. 230-233, April.

Mackay, D., W.Y. Shiu, A. Maijanen, and S. Feenstra. 1991. Dissolution of Non-Aqueous Phase Liquids in Groundwater. *Journal of Contaminant Hydrology.* v. 8, pp. 23-42.

Malone, D.R., C.M. Kao, and R.C. Borden, Dissolution and Bioremediation of Nonaqueous Phase Hydrocarbons: Model Development and Laboratory Evaluation. 1993. *Water Resources Research.* v. 29, no. 7, pp. 2203-2213.

Mercer, J.W., and R.M. Cohen. 1990. A Review of Immiscible Fluids in the Subsurface: Properties, Models, Characterization and Remediation. *Journal of Contaminant Hydrology.* v.6, pp. 107-163.

Miller, C.T., M.M. Poirier-McNeill, and A.S. Mayer. 1990. Dissolution of Trapped Nonaqueous Phase Liquids: Mass Transfer Characteristics. *Water Resources Research.* v. 26, no. 11, pp. 2783-2796.

Powers, S.E., C.O. Loureiro, L.M. Abriola, and W.J. Weber, Jr. 1991. Theoretical Study of the Significance of Nonequilibrium Dissolution of Nonaqueous Phase Liquids in Subsurface Systems. *Water Resources Research.* v. 27, no. 4, pp. 463-477.

Powers, S.E., L.M. Abriola, and W.J. Weber, Jr. 1992. An Experimental Investigation of Nonaqueous Phase Liquid Dissolution in Saturated Subsurface Systems: Steady State Mass Transfer Rates. *Water Resources Research.* v. 28, no. 10, pp. 2691-2705.

Powers, S.E., L.M. Abriola, and W.J. Weber, Jr. 1994. An Experimental Investigation of Nonaqueous Phase Liquid Dissolution in Saturated Subsurface Systems: Transient Mass Transfer Rates. *Water Resources Research.* v. 30, no. 4, pp. 321-332.

Smith, R.M. 1994. Effect of High Velocity Flow on the Dissolution of Contaminants From Residual Non-Aqueous Phase Liquids in Laboratory Column Studies. M.S. thesis, North Carolina State University. 117 pp.

Walas, S.M. 1985. *Phase Equilibrium in Chemical Engineering.* Butterworth Publishers, Boston, Massachusetts.

Wise, W.R., G.C. Robinson, and P.B. Bedient. 1990. Modeling Contaminant Leaching from Residual Oil. *Proceedings of the Petroleum Hydrocarbons and Organic Chemicals in Groundwater: Prevention, Detection, and Restoration.* Houston, Texas, October 31-November 2, 1990.

Wise, W.R., G.C. Robinson, and P.B. Bedient. 1992. Chromatographic Evidence for Nonlinear Partitioning of Aromatic Compounds Between Petroleum and Water. *Ground Water.* v. 30, no. 6, pp. 936-944.

**Distribution-Based Evaluation and Assessment of Mission Readiness
for the Evaluation of Personnel Training**

**David J. Woehr
Associate Professor
Department of Psychology**

and

**Michael J. Miller
Department of Psychology**

**Texas A&M University
College Station, TX 77843-4235**

**Final Report for:
Summer Faculty Research Program
Armstrong Laboratory**

**Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base
Washington, D.C.**

August 1995

**DISTRIBUTION-BASED EVALUATION AND ASSESSMENT
OF MISSION READINESS FOR THE EVALUATION
OF PERSONNEL TRAINING**

David J. Woehr
Associate Professor
Department of Psychology
and
Michael J. Miller
Department of Psychology

Texas A&M University

Abstract

The present paper briefly summarizes a research project focusing on ways to improve the usefulness of organization level outcome measures of unit readiness/effectiveness for the evaluation of personnel training interventions. Toward this goal, a measurement approach using organization level outcome measures is presented. It is suggested that an adaptation of this approach has the potential to improve the utility of organization level criterion measures. In addition, potential sources of data for the evaluation of training interventions at the organization level are identified and evaluated.

**DISTRIBUTION-BASED EVALUATION AND ASSESSMENT OF MISSION READINESS
FOR THE EVALUATION OF PERSONNEL TRAINING**

David J. Woehr

and

Michael J. Miller

Introduction

A vital concern for the Air Force is the maintenance of mission capability and readiness. A crucial mechanism for the maintenance of mission readiness is personnel training. There is little if any dispute that effective personnel training serves to enhance the effectiveness and capability of the Air Force in general. This fact is reflected in the overwhelming scope of training conducted throughout the Air Force and the tremendous amount of time and resources committed to the training endeavor.

Of tremendous importance to the design, implementation, and revision of training throughout the Air Force, as with any organization, is the ability to evaluate the effectiveness of training interventions. Goldstein (1991) defines training evaluation as: "the systematic collection of descriptive and judgmental information necessary to make effective training decisions related to selection, adoption, value, and modification of various instructional activities." (p. 557) More specifically, the effective evaluation of any training intervention is crucial to informed decision making regarding the intervention. Central to effective training evaluation is the standard or criteria against which the training is evaluated. In addition, the comprehensive evaluation of training interventions mandates the use of multiple criterion measures. The impact of training interventions must be assessed at different levels (e.g., person, work group, organization). Unfortunately, organization level outcome measures are often dismissed as criterion measures due to contamination by extraneous aspects of the work environment. Despite this limitation, the use of these measures is extremely important for demonstrating the utility of training interventions.

Organization Level Criterion Measures

Organization level outcome measures represent global indices of job effectiveness. They typically include results-oriented measures such as quality control indices, productivity or maintenance indices, promotion rate, salary progression or level, and turnover rates. The value of these measures as a standard for training evaluation is somewhat controversial. Two schools of thought can be found in the literature with respect to ways of conceptualizing the criterion construct. One school of thought emphasizes a conceptualization of performance as reflected in overt individual behaviors (e.g., Campbell, et al. 1970; Borman, 1983). This view focuses on the identification of behavioral regularities important to organizational functioning. The other school of thought focuses on outcomes. This view emphasizes the importance of outcomes and results to organizational functioning. Recent theories of the criterion construct, however, have begun to recognize the inextricable relationship between job behaviors and outcomes. Along these lines Binning and Barrett (1989) argue: "... optimal description of the performance domain for a given job requires careful and complete delineation of valued outcomes and the accompanying requisite behaviors" (p. 486).

Problems with Outcome-Based Criterion Measures

The detailed conceptual delineation of the relationship between job performance and outcomes is especially relevant to training evaluation. An important direction for future research is a focus on behavior/outcome linkages and generating empirical support for these linkages. Unfortunately, the operationalization of specific outcome measures generates somewhat of a dilemma for training evaluation. On the one hand, the ultimate value of training lies in its ability to impact outcomes of value to the organization. Outcome measures (eg., productivity levels, turnover rates, error rates, etc.) at both individual and aggregate levels would appear to be the ultimate criterion of interest for evaluating training interventions. On the other hand, these measures suffer from a number of problems that limit their usefulness as a standard against which to judge the impact of training.

First and foremost among these problems is the fact that these measures are typically contaminated to an undetermined extent by sources of variance over which the individual has no control. Specifically, the measured outcome is to some extent determined by factors other than

individual performance. A second problem with outcome measures is that they are not based on a common metric. Outcome measures are often unique to particular units within an organization and thus are difficult to interpret and compare across organizational work groups or divisions. Additionally, the lack of a common metric typically precludes the meaningful aggregation of performance information across organizational units. A third problem is that these measures only provide an indication of outcome as opposed to the process underlying the outcome. Thus these measures provide little, if any, information about the nature of performance. Finally, the traditional use of outcome measures offers little, if any, means of assessing measurement quality (i.e., how good are the measurements obtained with these measures).

Thus, although regularly collected and typically readily available, as a criterion against which to judge the impact of various training interventions in organizations, outcome measures have not proven as useful as criteria which are defined in terms of individual behavior. Despite this, however, the use of these measures is extremely important for demonstrating the ultimate utility of training interventions. Consequently, an important goal with respect to training evaluation is the development of ways to improve the utility of organization level criterion measures. Toward this end, a specific measurement approach to outcome-based criterion measures is presented below.

A Distributional Approach to Criterion Measurement

The measurement approach presented here extends the system for assessing individual performance developed by Kane (1986) to outcome level criteria measurement. It is believed that this approach may offer a partial solution to the problems associated with outcome measures. The original system presented by Kane (1986), labeled Performance Distribution Assessment (PDA), is based on the distributional measurement model postulated by Kane and Lawler (1979). An important characteristic of this model is a focus on the range of performance observed. Specifically, the model stipulates that not only is the level of performance important, but the fluctuation or variance in performance must also be considered. For example, two individuals may both be appropriately characterized as "average performers"; however, if one is consistently average and the other alternates between very poor and very good, very different

pictures emerge with respect to the individuals' performance. Thus performance measurement must assess the range of performance over time. Specifically, performance is defined in terms of the outcomes of job functions that are carried out on multiple occasions within a specified time span (i.e., iterated job functions). It is expected that, due to varying levels of individual ability and motivation as well as varying levels of external constraints, these outcomes will reflect different levels of effectiveness. Performance can subsequently be represented in terms of the frequency at which various outcome levels occurred within a given time span.

Another important characteristic of the PDA approach is that it incorporates a relativistic scaling of performance information. More specifically, performance is expressed as a ratio of actual performance (as reflected in the performance distribution generated) to a maximum feasible performance distribution. This maximum feasible distribution reflects the highest level of performance attainable given the constraints under which the work occurs. This scaling process serves to express performance in terms of a relative range of potential performance. Thus, the method allows for quantifiably excluding from consideration in the evaluation of performance the range of performance that is attributable to circumstances beyond the performer's control.

The representation of performance in distributional form along with relativistic scaling has several important advantages. First, it allows for a consideration of performance variability as well as average levels of performance. Thus it allows for an assessment of the consistency of performance and the extent to which negatively valued outcomes are avoided. In this way more information is provided regarding the idiosyncratic nature of individual performance. Second the relativistic scaling process advocated by the PDA process produces measures of the effectiveness of performance on relativized 0-100% scales with common zero and common upper limits of 100%. Thus any given percentage level remains constant in its meaning regardless of the job, division, or even the organizational level in which it occurs. At the same time, the particular outcome measures used to assess performance may be individualized to meet situational demands and organizational constraints. Specifically, if positions have appreciably different content and extraneous-constraint conditions, measures can be scaled to account for these differences.

The PDA approach was originally advocated as method for enhancing performance ratings. Specifically, it was formulated to incorporate subjective estimates of individual performance outcome frequencies (i.e., supervisory ratings of the frequency at which individuals performed at a particular level). However, its focus on the frequency of particular performance outcomes make it particularly amenable to use with more objective outcome measures. Thus, the application of this methodology to the measurement of organizational outcomes using iterative operational measures appears to be a fruitful avenue for research and may serve to increase the utility of these measures in the training evaluation process.

Adaptation of the PDA Approach for Outcome Level Measures

As noted above, the PDA system appears to be well suited for the measurement and scaling of operational criterion measures. For purposes of illustration, Table 1 presents hypothetical evaluation data presented in PDA format. In Table 1 the outcome range represents 5 equidistant steps between the highest possible performance outcome (listed as 95 in the Table) and the lowest acceptable performance level (listed as 75 in the Table). These values are based on supervisory or SME estimates. The utility weights represent the utility or value to the organization of performance at each of the 5 levels. Once again these values are based on SME estimates. Supervisors/SMEs also provide estimates for the "maximum" feasible distribution of performance as well as estimates of the "actual" performance level distribution of the individual being evaluated. Other values in the Table are calculated from the supervisor estimates. Here, it should again be noted that the PDA process was originally intended as a subjective performance rating system. Thus many of the data distribution utilized are based on supervisor or SME estimates of both subordinate performance and minimum and maximum performance levels. An adaptation of PDA for use with operational measures does away with these subjective estimates and replaces them with actual frequencies based on archival records of the operational measures. Also, the maximum feasible distribution is replaced with a "benchmark" distribution. This "benchmark" distribution may represent either an estimated ideal distribution of performance or the actual performance distribution of a comparison unit (i.e., an earlier time frame or another work unit). This revised approach is labeled here as Distribution-Based Evaluation and Assessment of Mission Readiness (DEAMR). This approach extends the

beneficial characteristics of relative distribution based performance assessment to organization level outcome measures. More specifically, characteristics of the DEAMR process include:

1. Performance measurement is relativistic. Outcome measures are scaled relative to maximum possible and minimum acceptable performance levels. Performance distributions are relative to some "benchmark" distribution. Thus, measurement considers the extraneous factors that may influence outcome measures.
2. Performance measurement is based on common metric. All measures are expressed in terms of percentages and thus have minimum and maximum points.
3. Multiple measures of performance are provided; performance is described in terms of mean level, consistency, and negative range avoidance. These multiple measures provide more information about the nature of performance and performance problems.

Another important characteristic of the DEAMR system is that it is easily automated. Relatively little data is required in order to calculate the distributional parameters. This data required includes the highest possible and lowest acceptable performance level, an estimate of the utility weight associated with the lowest acceptable performance level, and the actual frequency of performance outcomes at each of the performance levels. Table 2 represents the output of a spreadsheet based program specifically designed to provide distributional performance information. The highlighted boxes indicate where data must be input into the program. Performance distribution information is then automatically calculated and displayed both numerically and graphically.

Illustration of the DEAMR Process with Actual Data

An important aspect of the present study is the identification of potential data sources as well as the development of a system for improving the utility of outcome measures. Toward this end several existing Air Force data sources were identified and examined within the DEAMR framework.

CAMS Data

An initial potential data source examined was based on jet aircraft maintenance data gathered through the Core Automated Maintenance System (CAMS). This data provides the advantages of being regularly and systematically collected as well as being readily available in a usable format. The data examined in the present study was an archival six month sample (April - September 1990) of F15 maintenance data. CAMS data represents a record of aircraft maintenance activities performed. Table 2 presents an example of the output of this program following the DEAMR methodology. The primary CAMS based measure used in the DEAMR analysis is time to completion on various job tasks. Time to completion is based on the difference between the start and stop times for the task. Table 3 presents the DEAMR output for the "Benchcheck and Repair" action taken category derived from the CAMS data. Time to completion for job functions in this category ranged from .5 to 8. Distributional characteristics of performance in this category are presented in Table 3.

Limitations of CAMS Data

DEAMR analysis of the CAMS based time to completion data revealed a number of potential limitations of this data source. Some of these limitations were specific to the sample of data available. These included:

1. The data set was not collected or developed for this application (i.e., the 6 month sample had been culled from the larger data source for another purpose).
2. The number of cases available for analysis was quite small, thus making it difficult to make generalizations.
3. The units of comparison used were fairly broad (i.e., performing work center) making determination of appropriate benchmark distributions difficult.
4. The action category breakdowns are overly broad in that they do not consider on which equipment or system the action is performed.

Other limitations of the CAMS data resulted more generally from the nature of the data itself. These limitations represent a more serious threat to the usefulness of the CAMS based measures and include:

1. The CAMS-based measure is based purely on maintenance data and thus does not take into account unit characteristics such as manpower, mission requirements, etc.
2. The CAMS data is contaminated to an unknown extent by several factors including: the individual performing the work may not enter the data; standard or benchmarked time estimates may be entered instead of actual times; and, the measures are very heavily equipment driven.

Current Data Options

Given the limitations identified with the CAMS data, work was undertaken to locate other, more up to date, potential data sources. Two other current data sources were identified.

Mission Capability Data

A review of operational measures routinely collected at the 56th fighter wing at Luke AFB, Az. revealed one potentially valuable data source. This data represents a combination of CAMS- based maintenance data and unit mission characteristic data. This combination of data overcomes the major limitation associated with using CAMS data alone. More specifically, both equipment and unit mission and manpower characteristics are considered. Example measures include fully mission capable rate (FMC), man hours per flying hours, air and ground abort rates, etc. A listing of the actual data collected is presented in Table 4. The data presented in Table 4 is used to compute direct indices of mission capability/readiness (presented in Table 5). The indices presented in Table 5 have many desirable characteristics with respect to use with the DEAMR system. These characteristics include:

1. The measures are regularly and systematically collected.
2. It appears that these indices are both required by and reported to MAJCOM. Thus it is likely that these measures are available Air Force wide.
3. The mission capable/readiness indices reflect both equipment, mission, and manpower characteristics.
4. The indices are easily aggregated from the individual unit level to higher levels of the organization (wing, command, etc.).

5. The indices reflect multiple measures of performance within a specified time span (iterated job function) and thus are readily amenable to the DEAMR system.

These characteristics suggests further examination of these indices as criterion measures for training evaluation is warranted.

Quality Assurance Program Data

Another potentially valuable data set for use with the DEAMR system is quality assurance program data. Initial inquiry revealed that the 56th fighter wing has a wing based quality support office. This office regularly collects and disseminates quality control data relevant to fighter aircraft maintenance. This data is largely based on results of scheduled quality assurance inspections. Two indices appear to be particularly relevant for training evaluation. The first is quality assurance inspection pass rates. The second is the number of discrepancies per inspection. Discrepancies represent negative findings during quality assurance inspections. Both inspection pass rates and discrepancies are grouped by TEC code indicating the aircraft system or maintenance activity to which they apply. This data already provides valuable feedback to the training function and thus further examination of these measures as criterion measures for training evaluation is warranted.

Using DEAMR for Training Evaluation

Both the mission capability/readiness indices and quality assurance data represent viable potential criterion measures for the evaluation of training effectiveness at the outcome level. Further both data sources meet the requirements for use with the DEAMR system. Thus it is possible to rescale the data in distributional form. The DEAMR format could then be used to evaluate the effectiveness of specific training interventions.

While the data sources identified generally provide a potentially valuable source of information with respect to training evaluation, it is important to consider the further refinement of the data. More specifically, it would be beneficial to establish a pool of potential indices most relevant to specific training interventions to be evaluated (eg., maintenance indices for maintenance technician training). Here it is important to identify and evaluate key measures

from the larger pool of potential measures. The focus of this measure evaluation would be to identify indices that are: a) important to unit effectiveness, b) frequently and reliably collected, c) sensitive to individual performance, d) relatively insensitive to system variables, and e) relevant to training intervention. Information about outcome indices may be obtained through either SME workshops or through structured questionnaires. SME's would be used to provide information about each potential measure (eg., the relative importance of each measure, sensitivity to individual performance) as well as information relevant to the DEAMR process (eg., item utility weights, optimal possible and minimal acceptable levels, etc.). It is only through such systematic examination of the measures available that detailed conceptualizations of the linkages between individual performance and organizational outcome can be established. Ultimately, the effective use of outcome measures for the evaluation of personnel training depends on the delineation of these linkages.

References

Binning, J.F., & Barrett, G.V. (1989). Validity of personnel decisions: A conceptual analysis of the inferential and evidential bases. Journal of Applied Psychology, 74, 478-494.

Campbell, J.P., Dunnette, M.D., Lawler, E.E., & Weick, K.E. (1970). Managerial Behavior, Performance, and Effectiveness. New York: McGraw-Hill.

Goldstein, I.L. (1991). Training in work organizations. In M.D. Dunnette & L.M. Hough (Eds.) Handbook of Industrial and Organizational Psychology, vol. 2, (pp. 507-620). Palo Alto: Consulting Psychologists Press.

Borman, W.C. (1983). Implications of personality theory and research for the rating of work performance in organizations. In F. Landy, S. Zedeck, & J. Cleveland (Eds.) Performance measurement and theory (pp. 127-172). Hillsdale, NJ: Erlbaum.

Kane, J.S. (1986). Performance distribution assessment. In R. Berk (Ed.), Performance Assessment: Methods and Applications, (pp. 237-273). Baltimore, MD: Johns Hopkins University Press.

Kane, J.S., & Lawler, E.E. (1979). Performance appraisal effectiveness: Its assessment and determinants. In B. Staw (Ed.) Research in Organizational Behavior, Vol 1. Greenwich, CT.: JAI Press.

Table 1: Hypothetical Performance Distribution Data

Levels	1	2	3	4	5		Mean (SD)	Negative Range Score
Outcome Range	75	80	85	90	95		----	----
Utility Wts	-150	-87.5	-25	37.5	100		----	----
Maximum	0	.10	.15	.15	.60		53.13 (65.18)	-12.50
Actual	0	.10	.20	.30	.40		37.50 (62.50)	-13.75
Most Consistent	0	0	.20	.60	.20		37.50 (39.53)	- 5.00
Least Consistent	.067	.066	.267	0	.60		37.50 (82.30)	-22.50

EFFECTIVENESS OF THE ACTUAL DISTRIBUTION IS ESTABLISHED RELATIVE TO THE MAXIMUM FEASIBLE DISTRIBUTION - E.G.:

Relative Effectiveness of Mean of

$$\text{Actual Distribution} = (37.50/53.13) = 71\%$$

Relative Effectiveness

$$\text{of } \underline{\text{Consistency}} \text{ of } = (82.30 - 62.50) /$$

$$\text{Actual Distribution} = (82.30 - 39.53) = 46\%$$

Relative Effectiveness

of Negative Range

$$\text{Avoidance of Actual} = (-22.5 - (-13.75)) /$$

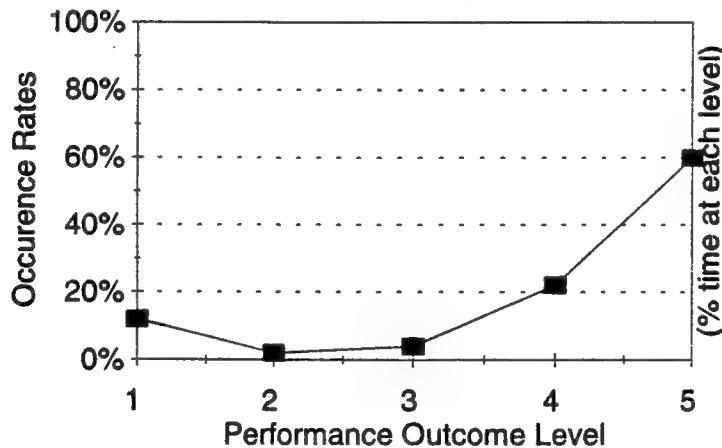
$$\text{Distribution} = (-22.5 - (-12.50)) = 88\%$$

Distribution Characteristics Calculations for:
 Unit: OVERALL
 Measure: Calibrate

Perf. Level	Perf. Range	Utility Weights	Performance Level Frequency	Performance Level Percentage	Distribution Characteristics	Utility Wt. Scale	PL Scale
1	8.00	-100.00	6.00	0.12	Mean =	58.00	4.16
2	6.13	-50.00	1.00	0.02	SD =	66.60	1.33
3	4.25	0.00	2.00	0.04	Skewness =	-1.56	-1.56
4	2.38	50.00	11.00	0.22	Kurtosis =	1.06	1.06
5	0.50	100.00	30.00	0.60	Negative Range Score =	-13.00	
Total Obs =		50.00	1.00				

Step 1 Enter min and max performance levels
 Step 2 Enter min. and max utility weights
 Step 3 Enter freqs. for each perf. level
 Perf level % are calculated
 Distributional are
 Characteristics calculated

Performance Outcome Distribution



Pg 15

Distribution Characteristics Calculations for:
 Unit: OVERALL
 Measure: Bench Check
 and repair

Perf. Level	Perf. Range	Utility Weights	Performance Level Frequency	Performance Level Percentage	Distribution Characteristics	Utility Wt. Scale	PL Scale
1	8.00	-100.00	3.00	0.15	Mean =	30.00	3.60
2	6.13	-50.00	1.00	0.05	SD =	65.95	1.32
3	4.25	0.00	2.00	0.10	Skewness =	-0.94	-0.94
4	2.38	50.00	9.00	0.45	Kurtosis =	-0.30	-0.30
5	0.50	100.00	5.00	0.25	Negative Range Score =	-17.50	
		Total Obs =	20.00	1.00			

Step 1 Enter min and max performance levels
 Step 2 Enter min. and max utility weights
 Step 3 Enter freqs. for each perf. level
 Perf level % are calculated
 Distributional are
 Characteristics calculated

Performance Outcome Distribution

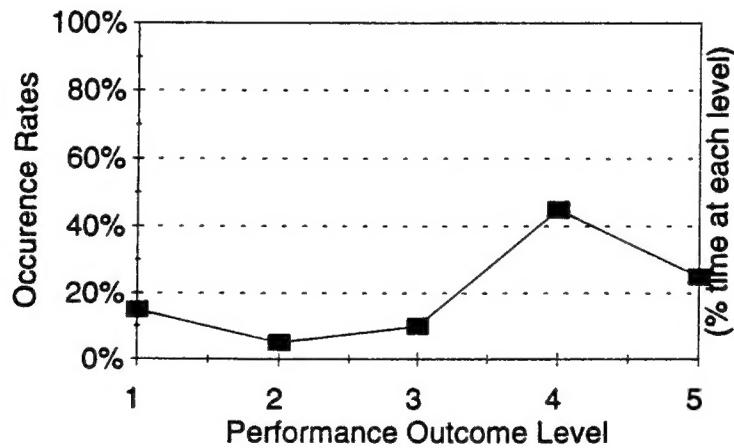


Table 4: Mission Capability Measures Collected

Data	Source
Authorized Aircraft	Wing Plans & Scheduling
Possessed Aircraft (aircraft available to flight, w possession codes of CA, CB,TF, or ZB0)	Wing Plans & Scheduling
Total Aircraft Possessed Hours (APH)	
Total Sorties Flown	
Total Man Hours	
Total Flying Hours	
Fully Mission Capable Hours (FMC)	CAMS # 459
Non Mission Capable Supply Hours (NMCS) " " " Maintenance (NMCM)	CAMS #459
" " " Both (NMCB)	CAMS #459
Partially Mission Capable Supply Hours (PMCS) " " " Maintenance (PMCM)	CAMS #459
# of 4/8/12 Hour Fixes	CAMS #460
# of Code 3 Breaks	CAMS #174
# of Ground Aborts	CAMS #174
# of Air Aborts	CAMS #174
# of Repeats (same malfunctin on next flight)	CAMS #174 & FS Recap Sheets
# of Recurs (same malfunction w/in next 3 flights)	CAMS #174 & FS Recap Sheets
# Aircraft Awaiting Parts (AWP)	
# Aircraft Awaiting Maintenance (AWM)	
Total Points Earned	
Total Points Scheduled	
Local Sorties Scheduled	
Weather Adds/Deletes	
Ferry/FCF Adds/Deletes	
Other Adds/Deletes	

Table 5: Mission Capability/Readiness Indices Calculated

<u>Measure:</u>	<u>Formula:</u>	<u>Calculated by:</u>
Avg. Possessed Aircraft (APA)	$\frac{\text{Total Actf Possessed Hrs}}{\text{Total Days in Month} \times 24}$	FS Analysts (daily)
Actual UTE Rate	$\frac{\text{Total Sorties Flown}}{\text{APA Aircraft}}$	FS Analysts (daily)
Manhours per Flying Hours	$\frac{\text{Manhours}}{\text{Total Flying Hours}}$	
Adjusted Sortie Schedule (Adj. Sortie Sched.)	$\begin{aligned} &\text{Local Sorties Scheduled} \\ &+ \text{Weather Adds} + \text{Ferry/FCF Adds} + \text{Other Adds} \\ &- \text{Weather Deletes} - \text{Sympathy Deletes} - \text{Other Deletes} \end{aligned}$	
Aircraft Scheduling Effectiveness	$\frac{\text{Adj. Sortie Sched.} - \text{Chargeable Deviations}}{\text{Adj. Sortie Sched.}} \times 100$	
Maintenance Plan Rate	$\frac{\text{Total Pts. Earned}}{\text{Total Pts. Scheduled}} \times 100$	
Mission Capable (MC) Rate	$\frac{\text{PMCM} + \text{PMCS} + \text{PMCB} + \text{FMC}}{\text{APH}}$	
Fully Mission Capable (FMC) Rate	$\frac{\text{FMC}}{\text{APH}} \times 100$	
Non Mission Capable (NMC) Rate	$\frac{\text{NMCS} + \text{NMCM} + \text{NMCB}}{\text{APH}} \times 100$	
Total Non Mission Capable Maintenance (TNMCM) Rate	$\frac{\text{NMCM} + \text{NMCB}}{\text{APH}} \times 100$	
Total Non Mission Capable Supply (TNMCM) Rate	$\frac{\text{NMCS} + \text{NMCB}}{\text{APH}} \times 100$	
Non Mission Capable Both (NMCB) Rate	$\frac{\text{NMCB}}{\text{APH}} \times 100$	

Table 5: Mission Capability/Readiness Indices Calculated (cont.)

Measure:	Formula:	Calculated by:
Partially Mission Capable Both (PMCB) Rate	$\frac{\text{PMCB}}{\text{APH}} \times 100$	
Total Partially Mission Capable Maintanence (TPMCM) Rate	$\frac{\text{PMCB} + \text{PMCM}}{\text{PAH}} \times 100$	
Total Partially Mission Capable Supply (TPMCS) Rate	$\frac{\text{PMCB} + \text{PMCM}}{\text{A PH}} \times 100$	
Ground Abort Rate	$\frac{\text{Ground Aborts}}{(\text{LCL Sorties Fln} + \text{Grnd Aborts})} \times 100$	
Air Abort Rate	$\frac{\text{Air Aborts}}{(\text{LCL Sorties Fln} + \text{Air Aborts})} \times 100$	
Total Abort Rate	Grnd Abort Rate + Air Abort Rate	
Code 3 Break Rate	$\frac{\# \text{ Code 3 Breaks}}{\text{Total Sorties Flown}} \times 100$	
4/8/12 Hour Fix Rates	$\frac{\# \text{ of 4/8/12 Hour Fixes}}{\# \text{ Code 3 Breaks}} \times 100$	
Deferred Discrepancy Rate (repair which cannot be done w/in 5 days)	$\frac{\# \text{ of AWM/AWP}}{\# \text{ of possed aircraft}} \times 100$	
Repeat/ Recur Rate	$\frac{\# \text{ Repeats or } \# \text{ Recurs}}{\text{Local Sorties Flown}} \times 100$	